Vol. 61, No. 2 March, 1954

Psychological Review

THEODORE M. NEWCOMB, Editor
UNIVERSITY OF MICHIGAN
Lorraine Bouthilet, Managing Editor

CONTENTS

Three Dimensions of Emotion ..... HAROLD SCHLOSBERG 81

A Mathematical Model and an Electronic Model for Learning ..... L. BENJAMIN WYCKOFF, JR. 89

A Statistical Theory of the Phenomenon of Subception ..... DAVIS HOWES 98

Reinforcement and Extinction Phenomena ..... JACK L. MAATSCH 111

Hypothetical Constructs and Intervening Variables ..... ARTHUR GINSBERG 119

Some Views on Mathematical Models and Measurement Theory ..... C. H. COOMBS, H. RAIFFA, AND R. M. THRALL 132

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

CONSULTING EDITORS

SOLOMON ASCH
ROBERT BLAKE
STUART W. COOK
CLYDE COOMBS
LEON FESTINGER
W. R. GARNER
JAMES J. GIBSON
D. O. HEBB
HARRY HELSON
E. R. HILGARD
CARL I. HOVLAND
E. LOWELL KELLY
DAVID KRECH
ROBERT W. LEEPER
ROBERT B. MACLEOD
DAVID C. MCCLELLAND
G. A. MILLER
GARDNER MURPHY
OSCAR OESER
CARROLL C. PRATT
DAVID SHAKOW
RICHARD SOLOMON
ELIOT STELLAR
S. S. STEVENS
ERIC TRIST
EDWARD WALKER
ROBERT WHITE

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

$6.50 volume $1.25 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized Jan. 8, 1948

Copyright 1954 by the American Psychological Association, Inc.

VOL. 61, NO. 2 MARCH, 1954

THE PSYCHOLOGICAL REVIEW

THREE DIMENSIONS OF EMOTION[1]

HAROLD SCHLOSBERG

Brown University

All of you have had to face the problems in the general field of emotion, whether your interest was theoretical or practical. I think you will agree that the field is chaotic. When you try to organize it, perhaps for presentation in a course, you probably follow one of two obvious methods. You can admit that "emotion is only a chapter heading," to quote Madison Bentley; in this case you present a sort of smorgasbord of interesting and important facts, and then go on to clinical cases or to experiments on drives in white rats, depending on your inclination. Or, if you wish a more orderly presentation of the topic, you may build it around some of the many theoretical controversies that stud the history of the field. My preference is for the latter method, and after years of following it in class, I think I am beginning to get a satisfactory integration. Let me run over the major theories, and show how they come together.

[1] This paper combines the material of two addresses: (*a*) "Some Dimensions of Emotion." Presidential Address, Div. 3, APA, Washington, Sept., 1952. Here the emphasis was on the qualitative dimensions, as found in facial expression. This material, in compressed form, appears near the end of the present paper, for the technical details of the study have since appeared in print (12). (*b*) "The Intensitive Dimension of Emotion." Vice-presidential Address, Sect. I (Psychology), AAAS, St. Louis, Dec., 1952. Read before a joint session of Sections I and Q (Education).

The first controversy was between the James-Lange theory and that dictated by common sense. There seems to be little doubt that James hit upon an important truth, namely, that the responses one makes in an emotional situation are more than mere expressions of a mental state; to put a current term in James's mouth, *feedback* from skeletal and visceral responses is an important component of an emotion.

The next major controversy was between the James-Lange and Cannon-Bard theories. If we stop worrying about whether the alleged mental state, emotion, resides in the cortex or in the thalamus, much of this controversy is pointless. Indeed, we can combine the contributions of the two theories and say that the hypothalamus is the key integrating center for outgoing impulses, and also for the feedback impulses that James emphasized; in this sense, it may be the "center" for emotions as well as for drives. This brings up another theory, the motivational theory of emotions, but this theory didn't meet much resistance, for anyone with a feel for either the derivation of the word *emotion,* or for the analysis of behavior, must agree that emotion and drive are overlapping categories.

The most recent controversy was over the question of whether emotion is organizing or disorganizing. My answer to this question is, "Both." Love cer-

tainly disorganizes a student's study habits, but it does organize certain extracurricular pursuits! On the other hand, one of the most challenging tasks of the teacher is to arouse interest, a mild emotion, in apathetic students. Clearly one can't answer the general question; whether emotion is organizing or disorganizing depends on (*a*) the task under consideration, (*b*) the nature of the particular emotion, and (*c*) the strength of the emotion. Duffy (2) has long emphasized the last two points, at times suggesting that we stop talking about emotion, and substitute the direction of behavior, and the strength of behavior; for the latter she used the phrase "degree of energy mobilization." It is a troublesome term, for some people quibble about the meaning of *energy* in this context, and others worry about the implications of *mobilization*.

There are many other terms that may be used for this intensitive dimension. Cannon's concept of preparing for an emergency contains the same basic idea, but doesn't give us a good term. Perhaps the best name for the dimension is that used by Lindsley, "activation." His chapter in the recent handbook (9) gives us the outline of what he calls an activation theory of emotion. The term *activate* means a bit more than to make active; the dictionary tells us that it also means to make reactive. Activation would seem to be a very good name for what emotion does to us; the angry man overreacts to stimulation. Strong emotion thus represents one end of a continuum of activation; the other end, the condition of minimum activation, is found in the sleeping man who doesn't respond to stimulation. (If we wish to be accurate, we should put the state of zero activation at death, rather than sleep, for the sleeping man may respond to very strong stimuli. But psychologists don't study organisms at activation levels below that of deep sleep!)

To illustrate what we mean by the continuum of activation levels, let us start with a sleeping man, one near the zero level of activation. His cerebral cortex is relatively inactive, showing only slow bursts of electrical activity on the electroencephalograph. The muscles are relaxed and send few return impulses to the central nervous system. The sympathetic, or emergency, division of the autonomic nervous system is fairly inactive. As a result of this general condition, he doesn't respond to ordinary stimuli; he is unconscious.

Now let the alarm clock ring. It is a strong stimulus, and breaks through the high threshold. Gross muscular responses occur, and feed back impulses into the central nervous system. There is also autonomic discharge, and the resulting responses of muscle and gland lead to more feedback, probably through some interwoven pathways, the reticular substance. These impulses reach the hypothalamus, increasing its level of activity, and this center activates the cerebral cortex, as can be seen from brainwaves (9). In short, the individual is awake and responsive to stimulation. Perhaps I should have said, "more or less awake," for some individuals take a lot of time and activity, with resulting feedback, before the level of activation is high enough to permit anything but routine activities like dressing and lecturing!

Let us assume that our hero has reached an optimum level of activation by 10:00 A.M. He is alert, and responds efficiently to his environment. But now he finds that a book he needs is missing from his shelf. This frustration produces an increment in level of activation, perhaps not high enough at first to be dignified by the name of anger. But as he continues to search for the book the level of activation

builds up until he is "blind with rage" or "functionally decorticate" to use Darrow's term; he probably wouldn't find the book now if it were under his nose. We will leave our hero in the range of level of activation that is conventionally set off as *emotion,* but let us not forget that he started at the other end of the continuum, sleep. The tendency to consider emotion as a separate state, divorced from the rest of the continuum, may well be the reason we have made so little progress in the field.[2]

Now let us consider level of activation in a more critical fashion: How are we going to measure it? Level of activation is a construct, crude at present. It is somewhat like level of prosperity. Everyone says that our country is more prosperous now than it was during the great depression. Suppose you wish a more precise statement, and consult an economist. He will quibble a bit, but ultimately he will probably suggest the use of a composite index, preferably based on key items like commercial bank deposits and payrolls of several large industries. Our economist will warn you that this gives an index of *general* level of prosperity, and that the level may differ in specific regions or industries.

Similarly, level of activation must be a general index, at least for the present. What are the key processes that we can use? There are a host of them, traditionally listed as bodily expressions of emotions. Blood pressure, heart rate, breathing indices, and hand steadiness are typical. A very promising one is tension in skeletal muscles, preferably accessory ones, as the brow potentials recorded by Kennedy and Travis (7, 14), or the neck muscles, with their important role in posture (10). But perhaps the most widely used measure is electrical skin resistance.

The psychogalvanic reflex, or better, the galvanic skin response, has been studied by hundreds of investigators. Perhaps its chief attraction is its sensitivity in mirroring ideational activity, particularly of an emotional nature. But its very sensitivity is largely responsible for the continuing argument as to whether or not GSR is a measure of emotion, for the response may be evoked quite readily by any sudden and strong stimulus, as a loud noise or an electric shock. Fortunately, the argument largely disappears if we drop the idea that emotion is a special state; strong stimuli, preparation to make an effort, and significant ideas all have a common feature, a quick increase in level of activation. This was recognized by Landis and Hunt (8) many years ago, when they stressed the fact that GSR is associated with an increase in subjective tension. As a matter of fact, tension is probably the popular word that comes closest to level of activation.

Consideration of the physiology of the GSR also points to its value as an index of level of activation. Darrow (1) showed that the fall in skin resistance was associated with the secretory activity of the sweat glands, under control of the sympathetic system. The skin resistance thus serves as an index of sympathetic discharge, a key element in the activation mechanism. He also pointed out that it is best to place the electrodes on the palms or soles, since these areas are relatively independent of thermoregulatory sweating. Finally, he suggested the best units in which to measure the phenomenon; rather than ohms, the usual measure of resistance, he preferred mhos, the re-

[2] Lindsley ends his chapter with the sentence, "In short, the activation theory appears to account for the extremes, but leaves intermediate and mixed states relatively unexplained as yet" (9, p. 509). The present discussion is less conservative.

ciprocal unit which describes conductance. This recommendation was based on the linear relationship he observed between conductance in mhos and the rate of sweat secretion. Later he suggested the use of log conductance, but a number of recent studies show that the mho is an excellent unit; it is normally distributed, and independent of the original level of resistance.[3] Conductance has the further practical advantage that it runs in the right direction, for increase in conductance is associated with increased level of activation, increased tension. Parenthetically, it is unfortunate that more workers don't calibrate their instruments in mhos, for it is at least questionable to perform even such a simple statistical operation as averaging on a skewed measure like the ohm.

The vast bulk of research on electrical skin conductance has lost the forest for the trees. Preoccupation with transient changes, the PGR, has led to general neglect of the slow drifts in absolute level of conductance, despite the fact that the absolute level is the obvious correlate of general level of tension or activation. This neglect is partly due to the design of conventional apparatus, which is adjusted to balance out basic level of conductance so that the transient changes may be read directly from deflections of a needle. For direct measures of level of conductance a much less sensitive and less elaborate apparatus is more adequate and convenient. I use a 50-microampere panel-type meter in series with a pair of dime-sized silver–silver chloride–saline paste electrodes. The potential to run the circuit is usually one volt, obtained from a flashlight cell with the aid of a potential divider and calibrated resistor; with this voltage, the dial needle reads directly in micromhos, and the range is adequate for most subjects. The whole gadget can be assembled for about $25, and is as portable as a box of cigars. It takes five minutes to attach the electrodes and adjust the apparatus, but only five seconds to obtain and jot down successive conductance readings. Thus, a single experimenter can carry on an experiment such as reaction time, taking periodic readings of conductance throughout the session. Although I haven't tried it yet, there seems to be no reason why this apparatus shouldn't be used to follow mean conductance level of groups of ten or more individuals engaged in a common task. For example, a pair of simple and reliable electrodes could be attached to each member of a small audience and connected in parallel to one meter. Since conductances in parallel summate, one would merely have to divide the total conductance by N to get the mean. Of course, small fluctuations, as asynchronous PGR's, would balance out, but the method should be perfectly adequate to determine the general changes in level of activation during the various episodes of a play, for example.

Now let me describe a few applications, to show that conductance serves as a satisfactory measure of level of activation.[4] One of the most convincing experiments is that of Duffy and Lacey (3). They recorded skin conductance on subjects who were going through several cycles of a psychophysical task. Conductance showed a sharp increase at the beginning of the first series of tones to be judged, and slowly fell during the progress of the series and ensuing rest period. Conductance shot up again at the start of the next series, dropped off during the series,

[3] See (13) for references.

[4] Skin conductance is not an ideal index of general level of activation. But short of a compound index, conductance may well be the best available, assuming that reasonable care is taken to keep electrode contact and room temperature fairly constant.

and so on. This saw-tooth pattern continued throughout the session, showing that the subjects alerted themselves each time they started a task, and then gradually relaxed as they made progress. Further, the general level of the conductance pattern fell from series to series within each session, and from day to day; the subjects were gradually relaxing as they became more familiar with the general situation.

Schlosberg and Stanley (13) have obtained results in an extensive series of tasks and tests run in five cycles over a two-hour session on each of five successive days. In addition to confirming the Duffy and Lacey findings, these experimenters hoped to relate conductance to efficiency. They ran into difficulties in the latter respect, for there was some suggestion of a curvilinear relationship between conductance and efficiency. This is what one would expect from what we said earlier about level of activation, for it seems likely that there is an optimal level of activation for each type of task, and perhaps for each subject. For example, a moderate level of activation would seem optimal for playing chess, whereas a relatively high one would be best for sprinting. In either case, the subject would report that he was too sleepy to do well if he were below the optimal level, and too tense if he were above it.

Plausible as this idea sounds, it is hard to pin down (10), for a lot of measurements are needed on each subject before we can obtain a good curve relating his efficiency to his level of activation. Freeman (4) showed one way to do it. He took short series of reaction times and simultaneously recorded skin resistances at various times during the day, depending on diurnal variations to give a broad range of levels of activation. His results showed a very clear inverted **U** relationship, with minimum reaction times at a moderate level of conductance. His published data are a bit scanty to establish such an important generalization, so it seemed desirable to repeat the experiment. I set up a portable reaction timer, with a built-in device to vary the foreperiod so that it could be used conveniently at home. A student has taken a hundred sets of readings on herself, sampling all hours of the day from before breakfast to bedtime. Each session included (*a*) conductance, (*b*) 20 simple auditory reaction times, and (*c*) hand steadiness. She obtained beautiful inverted **U** relationships between both hand steadiness and simple auditory reaction time on the ordinate and skin conductance on the abscissa. Her optimal level of conductance for hand steadiness is a trifle higher than that for reaction time. She also ran some short series of sessions on five other subjects, and they seemed to give comparable curves. These results are encouraging, for they seem to open the way to much fruitful work in the fields of skill, efficiency, and fatigue.

But you may feel that I have gotten quite far away from my title, "Three Dimensions of Emotion." Of course I have been dealing with the level of activation continuum, but perhaps you would like some studies on the high level of activation that is traditionally called *emotion*. I don't have anything very specific to report here, for emotions are hard to produce in the laboratory. I do have another student working on the effects of electric shock on conductance during a reaction-time task, but he has been too tenderhearted in adjusting the strength of the shock. However, I can at least point to the familiar lie detector test as a practical application of level of activation. The peak-of-tension method depends on a gradual increase in level of activation as the critical question approaches, followed by a marked fall in tension after the crisis has passed. Most interroga-

tors prefer breathing and blood pressure as indices of tension, on the grounds that PGR is too sensitive (6). On the other hand, the PGR has been used successfully, and it seems probable that even better results would be obtained with a device designed to measure absolute level of conductance, rather than the quick swings.

The activation theory of emotions has one obvious failing: It deals only with the intensitive dimension, and takes no account of differentiation among the various emotions. This is true as long as we limit our consideration to *general* level of activation, forgetting the fact that different subsystems might vary more or less independently in different emotions. The situation is quite analogous to that in the closely related field of motivation. Both hunger and thirst will raise the general level of drive, as measured on an activity wheel, but each will also act selectively on an appropriate family of S-R units. Unfortunately, the analogy isn't complete, for we haven't yet found differentiated emotional patterns as clean-cut as are eating and drinking. There are a few hints of such differentiation among bodily changes in the various emotions, but we need much more research before they can be established.

However, there is one field within the topic of emotions where we have long been embarrassed by the excessive number of different patterns: This is the field of facial expressions. Frois-Wittmann (5) brought some order out of this chaotic field by working out the interrelations among various expressions, and Woodworth (15) contributed a six-step scale that helped a lot. In 1941 Schlosberg (11) used this scale for collecting data on a new series of pictures, and found evidence that the scale described a roughly circular surface. Inspection of pictures arranged around the scale suggested that the surface might be generated by two axes, pleasantness-unpleasantness, and attention-rejection. The next step was to try to get a better description of the surface in terms of these two axes. In 1952 Schlosberg (12) reported the results of several attempts to obtain independent ratings on a large number of posed expressions, using nine-point rating scales, one for each dimension. Pleasantness-unpleasantness offered no trouble, but there was considerable difficulty in explaining the attention-rejection dimension to the judges. We tried pointing out that rejection was the active opposite of attention, characterized by compressed lips, nostrils, and eyes, as though forcibly excluding the external object, but this effort met with only mediocre success. We finally hit upon the use of "anchors"—pictures selected from another series to illustrate the extremes of attention and rejection. This stabilized the ratings and enabled us to locate each expression on the roughly circular surface described by the two dimensions, P-U and A-R. These positions were validated by using them to predict Woodworth scale judgments of the same pictures. The predicted Woodworth scale positions correlated with the obtained ones with coefficients of .92, .94, and .96 in three independent experiments, utilizing two different sets of photos of posed facial expressions. Hence, we may feel considerable confidence that P-U and A-R are two basic dimensions of facial expressions in particular, and perhaps of emotions in general.

We can compare the two dimensions of the facial-expression surface to the blue-yellow and red-green axes of the color surface. This immediately suggests that there may be a third dimension, corresponding to visual brightness. The third dimension for facial expres-

sions might well be the intensitive one we considered earlier, level of activation. As a preliminary test of this possibility, we obtained ratings on the same pictures, this time using a rating scale that ran from sleep to tension. The results enable us to construct a crude three-dimensional figure, roughly comparable to the familiar Munsell color solid, and quite as irregular in shape (Fig. 1). The unpleasant pictures tend to show the highest levels of activation, with mirth at an intermediate level, while contempt, which combines pleasantness with rejection, has a rather low level of activation. The third dimension seems to clear up some expressions that are not separated by the original two axes; for example, grief, pain, and suffering all have the same P-U and A-R values, but grief is rated considerably below the other two expressions in level of activation.[5]

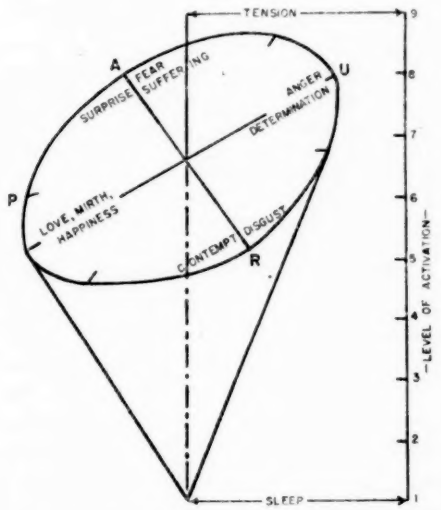

FIG. 1. A first approximation to the solid figure which represents the range of facial expressions. The emotions are placed correctly with respect to their maximum level of activation (sleep-tension) indicated on the ordinate. The top surface is sloped to show that anger and fear can reach higher levels of activation than can contempt. For a more accurate representation of the other two dimensions, pleasantness-unpleasantness and attention-rejection, see (12).

Much more work has to be done before we can be satisfied with the intensitive dimension of facial expressions. For one thing, we were working with a collection of pictures posed to represent emotions; this concentration on one end of the continuum introduces "series effects" in the ratings. We need a wider range of pictures, including low levels of activation such as sleep or listening to a dull lecture. Further, we should have actual skin conductance readings on the individuals, taken just before each picture was snapped. But these are projects for the future.

SUMMARY

The activation theory of emotion brings together many of the theories and facts of emotion, at least as far as the intensitive dimension is concerned. Instead of treating emotion as a special state, differing qualitatively from other states, the theory locates emotional behavior on a continuum that includes *all* behavior. This continuum, general level of activation, has its low end in sleep, its middle ranges in alert attention, and its high end in the strong emotions.

Any one of a number of physiological processes may be taken as an index of general level of activation, but electrical skin conductance has certain advantages for the purpose. It is sensitive, easy to measure, and varies in a manner consistent with expected changes in level of activation. It promises to be equally useful in work on skills and efficiency, as well as on emotions.

Neither skin conductance nor any other physiological measure of level of

[5] The three dimensions attempt to describe pictures of the *responses* called facial expressions. Knowledge of the situation which evoked a given expression will help the judge to interpret the expression, but such situational cues need have no part in the description of the response per se.

activation has yet given us much beyond the intensitive dimension. Further research may furnish such evidence, but for the present we may profitably turn to facial expression to find the qualitative dimensions along which emotion may vary. Here, we have good evidence that the whole range of expressions may be described rather well in terms of a roughly circular surface, whose axes are pleasantness-unpleasantness and attention-rejection. We have some idea how level of activation comes into this figure as a third dimension, but further research is needed here, too.

Thus, facial expressions and body changes supplement each other in giving us the dimensions along which emotions may vary.

REFERENCES

1. DARROW, C. W. The significance of skin resistance in the light of its relation to the amount of perspiration. *J. gen. Psychol.*, 1934, **11**, 451–452.
2. DUFFY, ELIZABETH. The concept of energy mobilization. *Psychol. Rev.*, 1951, **58**, 30–40.
3. DUFFY, ELIZABETH, & LACEY, O. L. Adaptation in energy mobilization: changes in general level of palmar skin conductance. *J. exp. Psychol.*, 1946, **36**, 437–452.
4. FREEMAN, G. L. The relationship between performance level and bodily activity level. *J. exp. Psychol.*, 1940, **26**, 602–608.
5. FROIS-WITTMANN, J. F. The judgment of facial expression. *J. exp. Psychol.*, 1930, **13**, 113–151.
6. INBAU, F. E. *Lie detection and criminal interrogation.* Baltimore: Williams & Wilkins, 1942.
7. KENNEDY, J. L., & TRAVIS, R. C. Prediction of speed of performance by muscle action potentials. *Science*, 1947, **105**, 410–411.
8. LANDIS, C., & HUNT, W. A. The conscious correlates of the galvanic skin response. *J. exp. Psychol.*, 1935, **18**, 505–529.
9. LINDSLEY, D. B. Emotion. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 473–516.
10. RYAN, T. A., COTTRELL, C. L., & BITTERMAN, M. E. Muscular tension as an index of effort: the effect of glare and other disturbances in visual work. *Amer. J. Psychol.*, 1950, **63**, 317–341.
11. SCHLOSBERG, H. A scale for the judgment of facial expressions. *J. exp. Psychol.*, 1941, **29**, 497–510.
12. SCHLOSBERG, H. The description of facial expressions in terms of two dimensions. *J. exp. Psychol.*, 1952, **44**, 229–237.
13. SCHLOSBERG, H., & STANLEY, W. C. A simple test of the normality of twenty-four distributions of electrical skin conductance. *Science*, 1953, **117**, 35–37.
14. TRAVIS, R. C., & KENNEDY, J. L. Prediction and control of alertness. III. Calibration of the alertness indicator and further results. *J. comp. physiol. Psychol.*, 1949, **42**, 45–57.
15. WOODWORTH, R. S. *Experimental psychology.* New York: Holt, 1938.

(Received March 22, 1953)

Psychological Review
Vol. 61, No. 2, 1954

A MATHEMATICAL MODEL AND AN ELECTRONIC MODEL FOR LEARNING[1]

L. BENJAMIN WYCKOFF, JR.

University of Wisconsin

In a previous report (11), the author outlined a quantitative theory of learning which would take into account the learning of observing responses. This theory was left incomplete largely because a quantitative statement of secondary reinforcement was required, and no suitable statement was available. The present development has grown out of an attempt to fill this gap by postulating a quantitative statement of secondary reinforcement. The resulting scheme has turned out to be surprisingly simple. The inclusion of this postulate, rather than complicating the system, has actually simplified it, since it enabled us to discard certain other postulates with no apparent loss in explanatory power.

In spite of the simplicity of the basic postulates of the system, some difficulty was encountered in deriving the implications for experiments involving a number of stimuli and responses. In particular, it was difficult to determine whether the postulate system would actually yield plausible predictions of observing response learning. To facilitate this process a specialized analogue computer was constructed to operate according to the postulates. The device constitutes a robot which can be confronted with learning problems. Its performance corresponds to the performance prescribed by the postulates of the mathematical model, thus establishing the implications of the theory for the particular learning problem employed.

The device described here makes use of relatively elementary electronic principles and is not intended as a contribution to the engineering aspects of computer design. It is of interest because of the learning theory embodied in it and as an illustration of the way in which the actual construction of a physical model may facilitate theory development in psychology.

The electronic model was confronted with a discrimination problem in which it was required to "learn" to select, by means of observing responses, those aspects of the situation which were relevant, in addition to learning the correct choice in a two-choice situation. By this method we were able to test whether the present theory would imply observing response learning. In the first attempts, serious shortcomings of the theory were forcefully demonstrated. The machine did not learn the correct observing response, and in fact learned the opposite of what was expected. It was possible to revise the theory and the model with these failures in mind. The model described here is the corrected model. The original failure will be considered in the discussion.

Both the mathematical and electronic models make use of time intervals as a critical variable. Originally the mathematical model was based on probabilities, but because of technical considerations a modified model using time intervals was substituted. A close cor-

[1] Many of the ideas expressed in this paper were stimulated by discussions with N. Guttman and C. J. Burke to whom the author is indebted. N. Guttman and W. S. Verplanck have developed a theory of reinforcement which parallels the present model with respect to qualitative aspects. C. J. Burke has developed a statistical model which treats observing response learning.

respondence between the time interval and probability models exists and will be considered.

THE MATHEMATICAL MODEL

The central postulate of the present theory derives from a suggestion made by Skinner (10, p. 246) that discriminative stimuli tend to exhibit secondary reinforcing properties. In other words, if some response has been strongly conditioned to a particular stimulus, then this stimulus will tend to strengthen a new reflex upon which it is made contingent. This notion was further elaborated by Notterman (8), Schoenfeld *et al.* (9), and Dinsmoor (2). Experimental evidence was obtained which suggested that establishment of a stimulus as a discriminative stimulus was necessary as well as sufficient for secondary reinforcement. Guttman and Verplanck [2] followed this line of reasoning in a slightly different direction and suggested that primary as well as secondary reinforcement may derive from the same principle. Their suggestion implies that primary reinforcing stimuli are effective because they are strong discriminative stimuli for a subsequent response which is sometimes called the consummatory response. The effects of deprivation on learning would be interpreted in terms of its effect on the strength of consummatory reflexes. The implications of this idea are far-reaching. However, further elaboration is beyond the scope of the present paper.

For our purposes it will be necessary to make a more explicit statement of the notion that discriminative stimuli exhibit secondary reinforcing properties. We note that the stimulus is considered as having a double role: first, its effect on immediately subsequent behavior which we call discriminative stimulus value or stimulus strength, and second, its effect on preceding S-R bonds which we will call reinforcing value. The present suggestion is that these two functions are directly related. In order to make a quantitative postulate on the basis of these considerations, we introduce the following definitions:

R = some class of "active" responses characterized by the fact that any member removes the subject from one stimulus to another.

Reflex = an S-R pair. No unconditioned connection is implied.

L = the total time that the stimulus of a reflex is present before the response occurs. This variable is not the same as the latency, in all cases, since the interval L need not be a continuous interval. For example, if R_1 failed to occur the first time S_1 was presented but occurred the second time, the value of L would be taken as the sum of the two intervals, whereas the latency would be the second interval only. The use of this variable will be indicated below.

V = *strength of a reflex* = $1/L$. This measure of reflex strength was selected partly because there exists a correspondence between it and *momentary probability* which has proved to be a convenient variable in theories such as Estes' statistical theory of learning (3). This correspondence will be discussed in more detail later.

V_t = *strength of a stimulus* = the strength of the reflex S-R_t, where R_t is the combined class of all active responses to a particular stimulus. V_t is called the strength of the stimulus because it refers to a particular stimulus but not to any particular response. It will be noted that $1/V_t$ is equal to the latency.

We are now in a position to present the quantitative rule of secondary reinforcement to be used in the present model. We postulate that if a response R_1 removes a subject from S_1 to S_2,

[2] GUTTMAN, N., & VERPLANCK, W. S. Personal communication.

the strength of the reflex S_1-R_1 is changed as follows:

$$\text{delta } V_1 = c(V_{t2} - V_1) \qquad \text{[Rule 1]}$$

where

V_1 = the strength of the reflex S_1-R_1,
V_{t2} = the strength of the stimulus S_2, and
c = a constant.

The equation states that V_1 is changed in the direction of V_{t2} by an amount proportional to the difference between them. If V_{t2} is greater than V_1, an increase in V_1 occurs; if it is less than V_1, a decrease occurs.

It will be seen that this rule is consistent with the previous qualitative statement of secondary reinforcement. In addition, it can be interpreted to include primary reinforcement if we assume the existence of at least one stimulus with a fixed maximum strength which will represent the primary reinforcing stimulus. If we insert a constant in place of V_{t2} in the above equation we obtain a function identical to the linear operator used by Bush and Mosteller (1). If we treat this equation as a differential equation we obtain functions formally identical to functions used by several theorists to describe simple learning (7, 3, 1).

In addition to the above rule of reinforcement we have found that it is necessary to the internal consistency of the model to include a second rule which relates to responses which fail to occur. The rule adopted states that if a stimulus S_1 is presented and the response R_1 fails to occur before some other response removes the subject from S_1 the reflex S_1-R_1 is changed as follows:

$$\text{delta } V_1 = -cV_1. \qquad \text{[Rule 2]}$$

This equation states that V_1 is decreased by a constant proportion of its value. The reason for including this rule will be discussed in more detail later.

These two rules are the only rules of change of reflex strength necessary to the present theory if we deal with "one-way" situations; that is to say, situations in which the subject never returns to the same stimulus twice during an experimental trial. An additional rule, which we will not elaborate, is required if retracing is allowed. Briefly, this rule provides for a decrement in any change in reflex strength if the subject returns to a stimulus soon after leaving it.

These postulates can be used to account for a variety of experimental findings in the sense that they yield behavior which is at least qualitatively the same as would be expected on the basis of experimental findings. This correspondence holds for findings of experiments on simple conditioning, extinction, delay of reinforcement, secondary reinforcement, and simple discrimination learning. The model also yields learning of observing responses under certain conditions and therefore can account for some kinds of "learning set," "concept formation," and stimulus generalization (11). It is expected that the model will also yield plausible predictions for intermittent reinforcement experiments and at least one kind of latent learning, although these cases have not been tested. The description of the electronic model and its operation will illustrate the way in which these postulates operate in the case of observing response learning in a two-choice situation.

THE ELECTRONIC MODEL

The electronic model consists of a group of variable time delay circuits, each representing a reflex. Each unit activates a relay when an input representing its stimulus has been connected for a cumulative total of L seconds.

The relay disconnects its own input circuit and connects the input for some other unit or units. The connections between the response of one unit and the stimulus of another are arranged by the operator so that all units and their connections form a *net* which represents some experimental situation. For example, a simple T maze might be represented by a net as follows:

$$\begin{array}{c} \text{-}R_L S_L \\ S_1\text{-}R_1 S_2 \\ \text{-}R_R S_R \end{array} .$$

This net is constructed of three reflex units S_1-R_1, S_2-R_L, and S_2-R_R. If the experimenter presented S_1, the S_1-R_1 time interval would begin. After L_1 seconds, R_1 would occur, removing S_1 and connecting S_2. At this time, both the S_2-R_L and S_2-R_R units would begin timing. If R_L occurred first, it would remove S_2 and connect S_L. Thus the machine's "choice" is determined by the time intervals of the units. On the next trial when S_2 was presented the two S_2 units would continue timing. Note that the S_2-R_R unit will already have some time registered from the preceding trial since R_R did not occur. Thus this unit will have a "head start." In this description we have ignored changes in reflex strength and simply illustrated the way in which the responses of the device are controlled by the delay circuits. We will now consider changes in reflex strength.

According to the mathematical model, changes in the strength of a reflex are determined by (*a*) the occurrence or nonoccurrence of the response before the stimulus is removed by another response, and (*b*) the strength of the following stimulus. The strength of the following stimulus is reflected in the latency of the following response, which we have noted is the inverse of the stimulus strength. For example, in the net described above changes in V_1 are to be determined by the latency (L_{t2}) of the subsequent response to S_2. If the response to S_2 is sufficiently prompt, V_1 is to be increased; otherwise it is to be decreased. The amount of change prescribed by Rule 1 was given as $c(V_{t2} - V_1)$ which is equivalent to $c(1/L_{t2} - V_1)$. For reasons of technical expediency, a negative growth function was substituted in the machine for the reciprocal function. This substitution provides a fair approximation over the critical range and greatly simplifies the circuit design.

If a response fails to occur before the stimulus is removed by the occurrence of some other response, V_1 is to be decreased by a constant proportion of its value. These changes in reflex strength and the corresponding changes in time delays are produced automatically by the electrical circuits. A value of .5 was used for the constant c in the present model. To save space, circuit diagrams and technical details will be omitted from the present report.[3] We are primarily concerned with the learning theory at this time. However, we may mention that the device employs relatively simple thyratron timer circuits. The strength of each reflex is represented by a voltage on a storage condenser, and this voltage is changed according to the postulates of the theory by appropriate relay connections.

Note that the strength of a reflex is affected by events which follow it in time, so that a unit which has just operated must be registered until the next

[3] To save printing costs, the circuit diagrams for the present model together with an eight-page discussion of the circuit's operation have been deposited with the American Documentation Institute. Order Document No. 4160 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C., remitting in advance $1.75 for 35 mm. microfilm or $2.50 for 6 by 8 in. photocopies. Make checks payable to Chief, Photoduplication Service, Library of Congress.

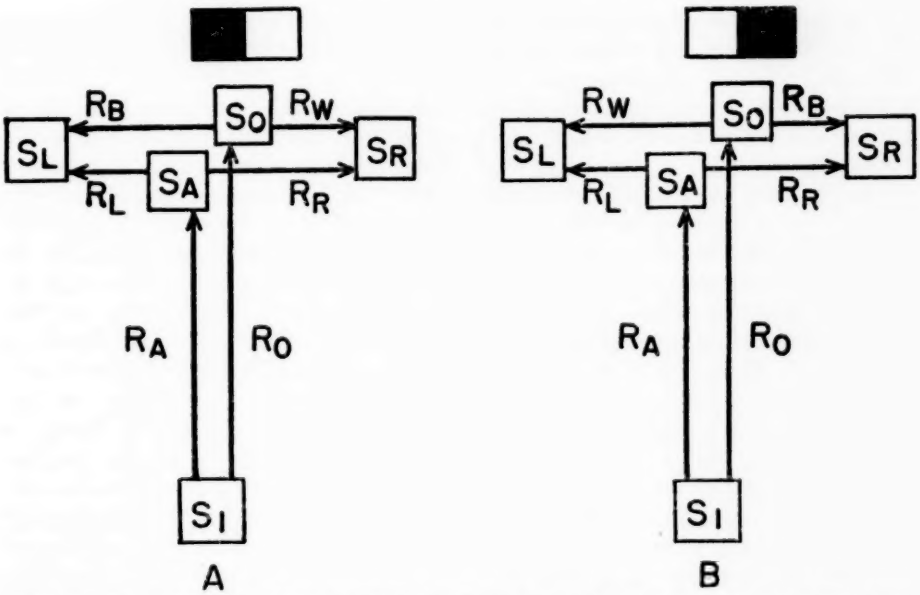

FIG. 1. The reflex net used to test for observing response learning. A. Six reflex units connected to represent a T maze with two stimulus cards at the choice point; a black card on the left and a white card on the right. B. The same T maze with the stimulus cards reversed.

response occurs. This system represents a compromise between a system such as Estes', which requires no "traces" or delayed effects, and other systems which require prolonged traces, or effects dependent on the final outcome of a trial.

All reflex units are identical in structure with the exception of a "primary reinforcing unit." This unit is a fixed interval timer with an interval of approximately .2 sec. The intervals of the other units may vary between approximately .2 and 80.0 sec. The units may be connected in any arbitrary net and may be attached to an array of lights on a panel representing some experimental situation. The operator may "run experiments" in much the same way that a rat experiment would be run, "placing the subject" at a starting point and connecting the primary reinforcing unit to some arbitrary position in the net. Latencies, rates, or relative frequencies of various responses can be recorded. In the following, we will describe a particular net which was set up to test for the learning of observing responses in a two-choice situation. This net is represented in Fig. 1. The diagram represents a T maze with two stimulus panels at the choice point. The experimenter may light either one so that it appears white while the other appears black.

From the starting box (S_1), two responses may occur, R_a or R_o. R_a may be thought of as the response of approaching the choice point with head down so that the stimulus panels are not visible. From S_a the subject may turn right (R_R) or left (R_L), but these responses will be independent of the position of the black and white panels. R_o, on the other hand, may be thought of as approaching the choice point with head up, so that the panels are visible. Responses from S_o are completely dependent on the position of the stimulus panels. They are specified as approach-

ing black (R_B) and approaching white (R_W). R_W takes the subject to the right goal box if the right panel is lighted and to the left goal box if the left panel is lighted. R_B is the opposite response.

The machine was connected to an array of lights representing the various positions in the net and the stimulus panels. The operator was then able to run experiments by switching on one of the stimulus lights, connecting the reinforcement unit to one of the goal box terminals, and presenting S_1. At this time, the S_1 panel light would appear. After an interval the light would "jump" to S_a or S_o, and then to one of the goal boxes. If it arrived at the correct goal box, the reinforcement unit would operate, turning off the goal box light after about .2 sec. If it arrived at the wrong goal box, the goal box light would simply stay on until the experimenter started the next trial.

Note that the reinforcement unit operates any time the subject arrives at the correct goal box, regardless of whether the initial response was R_a or R_o. Thus the subject may be reinforced even though it does not make the correct observing response. This corresponds to the fact that a subject in a discrimination experiment may make the correct turn even though it is not exposed to the discriminative stimuli. The experimenter places food in one of the goal boxes and the subject obtains the reinforcement when it arrives at the correct goal box even if it runs with its eyes closed. This general characteristic of situations involving observing responses is, in part, the source of the difficulties of predicting observing response learning from a quantitative theory.

Four kinds of discrimination problems can be tested in this net. In two of them "color" will be relevant, reinforcement always being presented for running toward one of the two colors. In the other two, "side" will be relevant, reinforcement always being presented for running to a particular side. Note that color discriminations can be solved only if the initial response is predominately R_o. Side discriminations depend on an initial response of R_a. We assume that the stimulus lights are reversed at random. Our interest is focused largely on whether the machine will learn to make R_a or R_o depending on whether reinforcement is consistently on one side, or consistently on one color. We will examine the results of one experiment in which the machine was confronted with six successive discrimination problems. The sequence of problems described in terms of the positive stimulus was as follows: white, black, white, right, left, black. Thus in the first problem, reinforcement was always connected to the goal box opposite the lighted stimulus panel regardless of which side this might be. In the fourth problem the reinforcement was always connected to the right-hand goal box, etc. This sequence provides several illustrations of reversals (white to black, right to left, etc.) and two illustrations of changes in the relevant aspect (white to right and left to black).

RESULTS

Figure 2 presents a learning curve for the entire sequence of six discrimination problems. The solid lines represent percentage of correct choices, that is, the percentage of response sequences terminating in the correct goal box. The dotted lines represent the percentage of correct observing responses. R_o was considered correct during color discriminations, R_a during side discriminations.

The curves in the first section of the graph show that both the correct choice and the correct observing response were learned while "white" was the positive

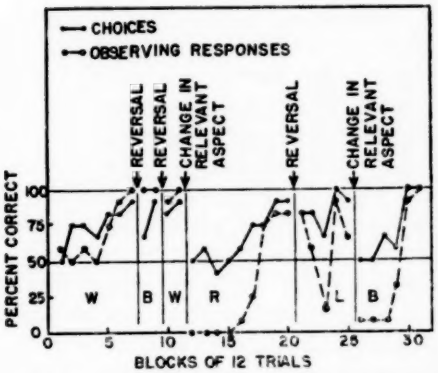

FIG. 2. Performance of the electronic model on six successive discrimination problems. The letters W, B, R, and L indicate the positive stimulus for each problem.

stimulus. In the second and third sections we note that reversed discriminations are learned very rapidly. The observing response which was learned during the first problem was still correct after the reversals and the percentage of correct observing responses remained high. On the fourth problem the relevant aspect was changed from color to side. The observing response which had been correct during the previous problems was now incorrect. We notice considerable retardation in learning on this problem. Learning of the correct observing response did not begin until after the fourth block of 12 trials. However, once this learning was started it proceeded fairly rapidly, reaching a reasonably high level within the next three blocks of 12 trials. In the fifth problem, following the reversal from right to left, we notice a temporary drop in the observing response curve, even though the same observing response was correct. This drop may be attributed to chance contingencies between incorrect observing responses and correct choices. Such contingencies may temporarily strengthen the incorrect observing response. Thus the learning of the reversed discrimination was also retarded in this case. The sixth problem shows the second change in relevant aspect and again shows retardation of learning at the beginning, followed by relatively rapid learning of the correct responses. In this case, the break in the curve was so sharp that it gave the impression of "insight" learning.

DISCUSSION

The present findings serve to illustrate the way in which a physical model may facilitate the development of a theory. By its use it was possible to show that the postulates of the theory predict the learning of observing responses and that, in this case, nothing outside the realm of this theory is necessary to account for observing response learning. Critics of reinforcement theory have sometimes argued that the facts of "attention" are incompatible with *S-R* reinforcement theories and that these facts compel us to make a basic revision in our theorizing. Hebb (6), for example, uses this argument to justify the introduction of a "central autonomous process." The present findings suggest that such critics underestimate the potential explanatory power of reinforcement theory.

The present development has also provided a clear illustration of the way in which a mechanical model may reveal the shortcomings of a theory. We mentioned earlier that the initial model failed to show observing response learning. A serious error in the original theory was not evident until the machine forced it to our attention. The original model operated on the basis of Rule 1 alone, so that no change in reflex strength was made for responses which failed to occur. When this model was tested on the observing response net (Fig. 1), it consistently learned not to make the correct observing response. The source of the difficulty, briefly stated, was that the strength of S_a remained too high during color discrimination learning. The combined

strength of R_R and R_L proved to be greater than the combined strength of R_W and R_B. On the basis of this observation the theory was re-examined. It became apparent that some decrement in the strength of responses which failed to occur would be necessary. It was possible to prove this necessity entirely apart from the problem of observing response learning. Rule 2 was added to correct this discrepancy, and further tests showed that the corrected model would yield observing response learning.

The variable V. The present model makes use of a somewhat novel measure of reflex strength, the variable V. The use of this variable enables us to maintain a certain degree of correspondence with a statistical model such as those proposed by Estes (3), Bush and Mosteller (1), and Estes and Burke (4), and at the same time allows for relatively simple mechanization of the model. The correspondence to the probability model will be evident if we examine the relationship between V and other measures of reflex strength. First, in a single-choice situation the reciprocal of V is the latency. In a two-choice situation, where two responses have strengths V_1 and V_2 the average latency will be $1/(V_1 + V_2)$. The relative frequency of R_1 will be $V_1/(V_1 + V_2)$, and of R_2 will be $V_2/(V_1 + V_2)$. In a free responding situation such as a Skinner box, the rate will be proportional to V. Parallel relationships hold if we use momentary probability (p) as a measure of reflex strength, where p is defined as the probability of occurrence of a particular response in a short time interval h. If we assume that p is the same at any moment during a given trial, the quantity p/h may be substituted in the above functions to obtain corresponding values for mean latencies and relative frequencies. (To obtain this correspondence it is necessary to assume that the time interval h is chosen sufficiently small so that the probability of two different responses occurring in the same interval is negligible.) It should be made clear that this correspondence holds for mean values only. The models do not correspond with respect to the distributions of these variables. This difference may or may not prove to be critical.

It is interesting to note that the present model, which was originally based on propositions growing out of reinforcement theory, shows some marked similarities to a contiguity theory, in that reinforcement is closely associated with removal of stimuli. However, if we examine the situation more closely we see that different stimuli are involved in the two cases. In Guthrie's system (5) the critical factor is the removal of the stimulus which was present at or before the time the response occurred, whereas in the present model reinforcement depends on removal of the stimulus which appears after the response is made, and we require that the stimulus be removed by the occurrence of a subsequent response.

SUMMARY

The present mathematical model was formulated with the objective of developing a quantitative model which would take into account the learning of observing responses. A quantitative postulate of secondary reinforcement plays an important role in this formulation. The postulate is based on a qualitatively expressed notion suggested by Skinner that secondary reinforcing properties of a stimulus are related to the discriminative stimulus value of the stimulus. By adopting quantitative definitions of reinforcement and "discriminative stimulus value" this proposition was readily translated into a quan-

titative postulate. This postulate forms the core of the mathematical model.

An electronic model was constructed to test whether the postulate system would yield plausible predictions of observing response learning. The device constitutes a robot which operates according to the postulates of the mathematical model. The electronic model was confronted with a discrimination problem in which it was required to select, by means of observing responses, those aspects of the stimulus situation which were relevant. The operation of the machine demonstrates that the theory will yield observing response learning. Some of the implications of the theory and the use of a physical model in theory construction are discussed.

REFERENCES

1. BUSH, R. R., & MOSTELLER, F. A. A mathematical model for learning. *Psychol. Rev.*, 1951, **58**, 313–323.
2. DINSMOOR, J. A. A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *J. exp. Psychol.*, 1950, **40**, 458–472.
3. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.
4. ESTES, W. K., & BURKE, C. J. A theory of stimulus variability in learning. *Psychol. Rev.*, 1953, **60**, 276–286.
5. GUTHRIE, E. R. *The psychology of learning.* New York: Harper, 1935.
6. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.
7. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
8. NOTTERMAN, J. M. A study of some relations among aperiodic reinforcement, discrimination training, and secondary reinforcement. *J. exp. Psychol.*, 1951, **41**, 161–169.
9. SCHOENFELD, W. N., ANTONITIS, J. J., & BERSH, P. J. A preliminary study of training conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1950, **40**, 40–45.
10. SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
11. WYCKOFF, L. B. The role of observing responses in discrimination learning: Part I. *Psychol. Rev.*, 1952, **59**, 431–442.

(Received April 1, 1953)

Psychological Review
Vol. 61, No. 2, 1954

A STATISTICAL THEORY OF THE PHENOMENON OF SUBCEPTION

DAVIS HOWES[1]

Aero Medical Laboratory, Wright-Patterson AFB, Ohio

Recent evidence reported by Lazarus and McCleary (**3**) leads to the conclusion that a conditioned autonomic response is a more sensitive indicator of the recognition of nonsense syllables than is the observer's symbolic report. For this phenomenon—sometimes described as autonomic discrimination without awareness—they have suggested the term *subception*. Its existence is of special interest because it invalidates a generalization that appears to hold for a great variety of perceptual experiments. This generalization may be formulated as follows: when an observer is given the task of discriminating among a set of stimuli, no measure of his success at that task is more sensitive than his symbolic (verbal) report. We shall refer to this as the *symbolic-report hypothesis*. Previous attempts to invalidate it have been numerous, but, as Lazarus and McCleary show in their report (**3**, pp. 113–116), satisfactory experimental controls were not observed in these earlier studies. Their own data, obtained in a very carefully designed experiment, thus provide the principal evidence against the hypothesis. It is the purpose of this paper to show that these data actually are not inconsistent with the symbolic-report hypothesis when certain statistical effects are taken into consideration.

The Experiment

The Lazarus and McCleary experiment consisted of two essential parts, a conditioning period and a test period. During the conditioning period, galvanic skin responses (GSR), elicited by electric shock, were conditioned to each of 5 nonsense syllables (shock syllables) exposed in a tachistoscope for 1 sec. Five other nonsense syllables (nonshock syllables) were exposed an equal number of times, but were not paired with shock. Observers were shown the list of 10 syllables and were told which 5 would be paired with shock and which 5 would not. In the test period, the 10 syllables were again exposed, but at 5 different durations, selected by pretest to range from durations too brief for accurate recognition to durations long enough to result in almost 100 per cent recognition. No shocks were given during the test period. The important data recorded during the test period were the GSR for a 5-sec. period following each exposure, and the observer's verbal report as to which of the 10 syllables had been exposed. Precautions taken by *E*s to control various artifacts are summarized in their report (**3**, pp. 116–118).

Three properties of the observer's reports were used to categorize the data. These, with the symbols used to represent them, are as follows: (*a*) the report was right (*R*) or wrong (*W*); (*b*) the report was a shock syllable (*S*) or a nonshock syllable (*N*); (*c*) the stimulus actually exposed was a shock syllable (subscript *s*) or a nonshock syllable (subscript *n*). The possible combinations of these properties define the six categories RS_s, WS_s, WN_s, RN_n, WS_n, WN_n. For each of these experimental categories Lazarus and McCleary report GSR

[1] Present address Wilmington, N. C.

in micromhos, averaged over all five exposure times.

A distinction needs to be drawn here between the *subception effect* and the *subception hypothesis*. The subception effect is defined by Lazarus and McCleary as the inequality

$$(WS_s + WN_s) > (WS_n + WN_n),$$

where the symbol for each category represents the mean GSR obtained under the conditions defining that category. This inequality was found to hold for all nine observers used in the Lazarus and McCleary experiment, and the mean difference in GSR could not be accounted for by chance variations ($t = 7.45$, $df = 8$). These results establish the existence of the *subception effect*. This effect leads Lazarus and McCleary to conclude (p. 118) that "subjects can make autonomic discriminations when they are unable to report conscious recognition" (i.e., correct symbolic report). Such a mechanism of "unconscious perception" can be called the *subception process*. No property of the experimental data other than the subception effect is offered by Lazarus and McCleary as evidence for a process of this type. It will be argued here that no subception process need be postulated, since the subception effect can be derived from the symbolic-report hypothesis.

THE STATISTICAL THEORY

Let us take for a starting point a statement of the symbolic-report hypothesis that is simply a negation of the subception hypothesis stated by Lazarus and McCleary: *GSR discrimination can occur only when the observer can report the exposed syllable correctly.* This formulation, which we shall refer to as the *classical* form of the symbolic-report hypothesis, is invalidated by the existence of the subception effect. But it does not take into account the statistical nature of discrimination. The theory proposed here considers the processes underlying recognition to be stochastic variables. This conception leads to the *statistical* form of the symbolic-report hypothesis: *at any specified moment, the GSR accompanying an observer's report is proportional to the probability that that report will be a shock syllable.* Like its classical counterpart, this formulation of the hypothesis implies that there is no autonomic discrimination without verbal discrimination, but it treats verbal report as a statistical process. A similar assumption has been proposed elsewhere (**1**) to account for the unusually long exposure times found to be required for the recognition of taboo words (**4**).

From the statistical form of the hypothesis it follows that any condition which alters the probability that a shock syllable will be reported will elicit a corresponding change in GSR. Tachistoscopic exposure of a shock syllable is such a condition. We shall follow the argument for a set of m syllables so exposed that they can be read when exposed for indefinitely long durations. Consider the effect of exposing a syllable for the limiting conditions of zero and of infinite duration. For zero duration of exposure all syllables will have the same probability of report if (as in the Lazarus and McCleary study) the m syllables are distributed equally over the experimental conditions. The probability of any exposed syllable is thus $1/m$. When the syllable is exposed for indefinitely long durations, however, the probability that it will be reported correctly approaches 1. These limiting conditions show that increasing the exposure time of a given syllable from zero to infinite

duration raises the probability that the syllable will be reported. All experimental evidence indicates that between these limiting conditions the relation between probability of correct verbal report and exposure duration is monotonic and nondecreasing (**2**, **5**, **6**, **7**). We then have the implication

$$[t_2(i) > t_1(i)] \subset [p_2(i) \geq p_1(i)], \quad [1]$$

where $p_2(i)$ is the probability that syllable i will be reported following its exposure for $t_2(i)$ seconds and $p_1(i)$ the probability of its report following exposure for $t_1(i)$ seconds. The symbolic-report hypothesis, in the statistical form from which the subception effect is to be derived, may be written

$$[p_2(i_s) \geq p_1(i_s)] \subset [c_2(i_s) \geq c_1(i_s)], \quad [2]$$

where i_s represents a *shock* syllable, $c_2(i_s)$ is the mean palmar conductance (GSR) measured when $p_2(i_s)$ is the probability of the observer's reporting i_s, and $c_1(i_s)$ is the mean GSR when $p_1(i_s)$ is the probability of his reporting i_s. From Equations 1 and 2 we have, by *modus ponens*,

$$[t_2(i_s) > t_1(i_s)] \subset [c_2(i_s) \geq c_1(i_s)]. \quad [3]$$

This conclusion states that, on the average, exposure of a shock syllable for a long duration will yield a larger GSR than exposure of the same syllable for a shorter duration.

Now of the two pairs of experimental categories that define the subception effect, $(WS_s + WN_s)$ includes only shock-syllable exposures while $(WS_n + WN_n)$ includes only exposures of nonshock syllables. Thus $t(i_s)$, the time for which shock syllables are exposed, must be zero for $(WS_n + WN_n)$ and greater than zero for $(WS_s + WN_s)$; and the inequality of GSR that defines the subception effect,

$$(WS_s + WN_s) > (WS_n + WN_n), \quad [4]$$

follows from Equation 3. Our argument to this point shows that the existence of the subception effect does not require that the symbolic-report hypothesis be abandoned, but only that the formulation of the hypothesis take into account the fact that recognition must be defined in statistical concepts.

FURTHER DEDUCTIONS FROM THE STATISTICAL THEORY

We can deduce more than the gross subception effect [4] from our statistical hypothesis. Lazarus and McCleary report mean GSR for all six of their experimental categories. A deduction can be drawn from our assumptions for each of the 15 possible comparisons between these means. To make the argument clear, however, we must analyze the concept of response probability a little more closely.

Corresponding to each of our m syllables we assume m characteristic underlying processes, each independent of the others. These processes will be regarded as analytic fictions, and no physiological referent need be ascribed to them. An observer will emit (report) the ith syllable only if the strength of the ith process is greater than the strength of the process underlying any of the other syllables. At this point we introduce our stochastic variable: we assume that the strength of the ith process, w_i, is determined stochastically, not uniquely, by the specification of experimental conditions. While we cannot fix w_i by experimental techniques, we can use such techniques to fix a distribution function $F(w_i)$. Suppose a set of conditions fixes the same distribution function for two syllables i and j. If the momentary strengths of the two syllables are observed upon an infinitely large number of occasions,

w_i will be the greater on half of these observations, w_j on the other half. Since the emission of a syllable may be considered to depend upon the integration of these momentary strengths over short periods of time, the probability (limit of relative frequency) of emission of syllable i will be equal to that of syllable j. We can therefore associate the condition of identical distribution functions with syllable probabilities of $1/m$ and with zero duration of exposure. The other limiting condition, where all values of the distribution function of the ith process are greater than the largest possible value of the process underlying any other syllable, will yield a probability of i equal to 1 and may therefore be associated with the condition of infinite exposure of i.

All we have done in the above argument is to identify an increase in the duration of exposure of the ith syllable with an increase in the probability that the ith process will exhibit greater strength than the process underlying any other syllable. The symbolic-report hypothesis can now be reformulated: at any specified moment, *the GSR accompanying an observer's report is proportional to the total strength of all processes underlying shock syllables.*

We must now assess the relative strengths of the shock-syllable processes in each of the six categories of the Lazarus and McCleary experiment. It must be remembered, however, that the report of the observer only tells us which of the various underlying processes is the strongest. Yet a shock syllable may contribute heavily to the GSR even if it is not reported by the observer, for its underlying process may be only slightly weaker than the strongest process. How can the relative strengths of these unreported syllables be analyzed? For the data reported by Lazarus and McCleary there is no way to obtain this information for a given exposure, but the assumptions we have made will suffice to determine averages. We shall restrict our analysis to the strongest and second-strongest underlying processes, since an analysis of third-strongest and still weaker processes would simply repeat the argument and reinforce the effects shown for the second-strongest processes.

Before proceeding to the argument, it will be convenient to summarize here the definitions and assumptions upon which the argument will be premised. *Definition:* Verbal report of the ith syllable means that, at the moment of report, the strength of the ith process is greater than the strength of any other process. *Assumption 1:* The probability that the ith process will have a strength w_i is a monotonic, nondecreasing function of the duration for which syllable i is exposed. *Assumption 2:* Exposure of any syllable other than the ith will have only a negligible effect upon the probability that the ith process will have a strength w_i. *Assumption 3* is the symbolic-report hypothesis. The argument will be carried through only for the special case in which all unexposed syllables are assumed to have identical distributions of process strength (i.e., are equally likely to be reported), although it can be extended to cover the general case in which each unexposed syllable is assigned a different distribution of strength.

Let the capital letters A through E represent the processes underlying the five shock syllables A' through E' of the Lazarus and McCleary experiment, and let the lower-case letters a through e represent the processes underlying the five nonshock syllables a' through e'. Consider the situation when a shock syllable, A', is exposed.

All observer reports placed in the RS_s category must then have A for their strongest underlying process. But by our assumptions, no one of the remaining nine processes is more likely than any other to be the second-strongest process. Consequently, nine equally probable combinations of strongest and second-strongest processes are possible for this category: (A, B), (A, C), (A, D), (A, E), (A, a), (A, b), (A, c), (A, d), (A, e). These alternatives are listed under RS_s in Table 1.

TABLE 1

POSSIBLE COMBINATIONS OF STRONGEST AND SECOND-STRONGEST SYLLABLE PROCESSES FOR EACH OF THE SIX CATEGORIES OF THE LAZARUS AND MCCLEARY EXPERIMENT

A′ Exposed	*a′* Exposed
RS_s	RN_s
A A A A A A A A A *B C D E a b c d e*	*a a a a a a a a a* *A B C D E b c d e*
WS_s	WS_n
B *B B B B B B B B* **A** *C D E a b c d e*	*A A A A* **A** *A A A A* *B C D E* **a** *b c d e*
C *C C C C C C C C* **A** *B D E a b c d e*	*B B B B* **B** *B B B B* *A C D E* **a** *b c d e*
D *D D D D D D D D* **A** *B C E a b c d e*	*C C C C* **C** *C C C C* *A B D E* **a** *b c d e*
E *E E E E E E E E* **A** *B C D a b c d e*	*D D D D* **D** *D D D D* *A B C E* **a** *b c d e*
	E E E E **E** *E E E E* *A B C D* **a** *b c d e*
WN_s	WN_n
a *a a a a a a a a* **A** *B C D E b c d e*	*b b b b b* **b** *b b b* *A B C D E* **a** *c d e*
b *b b b b b b b b* **A** *B C D E a c d e*	*c c c c c* **c** *c c c* *A B C D E* **a** *b d e*
c *c c c c c c c c* **A** *B C D E a b d e*	*d d d d d* **d** *d d d* *A B C D E* **a** *b c e*
d *d d d d d d d d* **A** *B C D E a b c e*	*e e e e e* **e** *e e e* *A B C D E* **a** *b c d*
e *e e e e e e e e* **A** *B C D E a b c d*	

Still assuming that A' is the syllable exposed, let us examine the WS_s category. Here the strongest process must by definition correspond to one of the unexposed shock syllables B', C', D', or E', and by our assumptions there is an equal probability that it will be each one. But the second-strongest process cannot be specified so simply. Take the case where B is the strongest process. Any of the nine remaining processes can be the second-strongest. But one of these is A, the process underlying the exposed syllable, and by our first assumption its average strength depends upon the duration for which A' has been exposed. Only for exposures approaching zero duration will the average strength of A be as small as the average strengths of the other eight processes. When A' is exposed and B' happens to be reported, then, the second-strongest process will probably be A. The remaining eight processes will have smaller probabilities of being second-strongest. The difference between the probability of A and the probability of each of the other processes will depend upon the duration for which A' has been exposed: for short durations the difference will be small; for long durations the probability that A is the second-strongest process will approach unity.

We have analyzed the WS_s category only for the case in which B' is the syllable actually reported. The same analysis obtains for the equally likely cases where C', D', or E' happens to be reported. In each case, A is the most likely process to be second-strongest. The 36 possible combinations of strongest and second-strongest processes for this category have been listed in Table 1. The combinations (B, A), (C, A), (D, A), and (E, A), which are

more probable than the others whenever A' has been exposed, have been printed in boldface.

Each of the other four categories used by Lazarus and McCleary has been analyzed in the same way, and all possible combinations of their strongest and second-strongest processes are listed in Table 1. The table is worked out for the cases where A' is the shock syllable exposed and a' is the nonshock syllable exposed. Combinations printed in the same typeface are of equal probability; those printed in boldface are the combinations whose probabilities are increased by exposure of A' or a'. For exposure durations approaching zero the probabilities of the boldface combinations approach those of the other combinations; for long exposures the probabilities of the boldface combinations become so great that the other combinations may be ignored.

Table 1 permits us to compare the shock-syllable strengths of any two categories for specified exposure durations. Now Lazarus and McCleary report only *average* GSR for all reports falling within each experimental category. In order to compare shock-syllable strengths averaged over all exposure durations, then, we must consider the extent to which each category will contain reports following long and short exposures. For exposures of zero duration, the probability that the "exposed" syllable will be reported will be no greater than the probability that any other syllable will be reported. The probability that a report will fall into the RS_s category if a shock syllable is exposed (or into the RN_n category if a nonshock syllable is exposed) will thus be only 1/10. For long exposures, however, the probability of the exposed syllable will be high, and all but a few of the observer's reports will fall into the RS_s (or RN_n) category. Most of the reports falling into the RS_s and RN_n categories will thus have been emitted following long exposures, while most of the reports falling into the other four categories will have been emitted following short exposures. This heterogeneity must be taken into account in comparing either RS_s or RN_n with any of the other categories.

We are now ready to determine which of each pair of categories will have the greater shock-syllable strength. In 8 of the 15 possible pairs of categories the strongest process of one category underlies a shock syllable while the strongest process for the other category underlies a nonshock syllable, and the proportion of long and short exposures in the two categories either will not differ appreciably or will reinforce the difference between their strongest processes. Inequalities of shock-syllable strength that will hold regardless of exposure duration can therefore be stated for these pairs of categories. Hence $RS_s > RN_n$; $RS_s > WN_s$; $RS_s > WN_n$; $WS_s > RN_n$; $WS_s > WN_s$; $WS_s > WN_n$; $WS_n > RN_n$; $WS_n > WN_n$. In these inequalities the symbol for each category represents the shock-syllable strength averaged over all responses falling into that category.

Table 1 contains another four pairs for which one category has greater shock-syllable strength for all durations of exposure except those approaching zero. At durations approaching zero the shock-syllable strengths of the compared categories approach equality. These also are pairs for which the proportion of long and short exposures either will not differ appreciably or will reinforce the difference apparent from Table 1. Since shock-syllable strengths must be averaged over all exposure dura-

tions, inequalities may also be stated for these pairs: $WS_s > WS_n$; $WN_s > WN_n$; $RS_s > WS_n$; $WN_s > RN_n$. The first two of these, it may be noticed, correspond to the subception effect reported by Lazarus and McCleary.

Consider next the RS_s and WS_s categories. Table 1 indicates that both categories will have equal shock-syllable strengths for exposure durations approaching zero. But for very long exposures almost all of the second-strongest processes in the WS_s category will underlie shock syllables, while 4/9 of the second-strongest processes in the RS_s category will underlie shock syllables. Reports following exposures of a specified long duration, then, should reflect a greater average shock-syllable strength if they fall into the WS_s category rather than the RS_s category, though the difference, being restricted to a portion of the second-strongest processes in the respective categories, will be small. The RS_s category, however, will contain up to 100 per cent of the reports following long exposures, while the WS_s category will include up to 4/5 of the short exposures. When shock-syllable strengths are averaged over all reports included in each category, then, the inequality will be $RS_s > WS_s$, for RS_s will be composed chiefly of large strengths derived from long exposures while WS_s will be composed chiefly of small strengths derived from short exposures.

The RN_n and WN_n categories also are identical for very brief exposures. Increasing exposure duration will not alter the shock-syllable strength of RN_n. The shock-syllable strength of WN_n, however, will *decrease* with increasing exposure duration, because the boldface combinations for this category include nonshock syllables as both strongest and second-strongest processes. On the average, therefore, reports following long exposures will reflect slightly smaller shock-syllable strengths if they fall into the WN_n category than if they fall into RN_n. But practically all the long exposures will fall into the RN_n category, while up to 4/5 of the short exposures will fall into the WN_n category. Pooling exposures of all durations will therefore reduce the average difference between these categories to a very small value.

Comparison of WS_n and WN_s involves a similar problem. The inequality $WS_n > WN_s$ will hold for short exposures. But increasing the exposure time will increase the average strength of the second-strongest processes of the boldface combinations shown for these categories in Table 1. As a consequence, the average difference between the strongest and second-strongest processes of each of these categories will decrease with increasing exposure, since the strongest process in each will underlie an unexposed syllable and hence will not gain appreciably in strength as a result of the exposure. With exposures for which nearly 100 per cent of reports are correct, the difference between the shock-syllable strengths of these categories will be negligible. These long exposures, for which the two categories approach equality, will produce greatest shock-syllable strengths and will therefore have the greatest weight in determining the average shock-syllable strength for each category.

In the last two comparisons, between RN_n and WN_n and between WN_s and WS_n, inequalities of shock-syllable strength can be shown only for exposure durations which produce small process strengths. For exposure durations producing maximum process strengths, however, the shock-

syllable strengths of each of these pairs of categories will approach equality. Since maximum process strengths will have the greatest weight in fixing the average shock-syllable strength of each category, the inequalities for these two pairs of categories will be seriously weakened. Except when the averages are taken over very large numbers of observations, these weak inequalities probably will be masked by the random variability that is a defined property of the syllable process and by the errors of measurement that will result from the effects of uncontrolled factors upon GSR. Now Lazarus and McCleary report the following numbers of observations per observer for the four categories involved: 24 (RN_s), 14 (WN_n), 13(WS_n), and 15 (WN_s). As the inequalities cannot be expected to hold reliably for such small numbers of observations, the relative shock-syllable strengths of these categories will be written $RN_n \approx WN_n$ and $WS_n \approx WN_s$, where the direction of each unreliable inequality is preserved in the order in which each pair of categories is written.

TABLE 2

TESTS OF THE DEDUCED INEQUALITIES

Column A lists the inequalities of shock-syllable strength. Column B shows the number of observers in the Lazarus and McCleary experiment (3) whose mean GSR confirms each inequality and the number whose mean GSR contradicts it.

The probability of obtaining 9 confirmations by chance is .002; 8 confirmations, .018; 7 confirmations, .070.

A	B	
	Yes	No
$RS_s > RN_n$	9	0
$RS_s > WN_s$	9	0
$RS_s > WN_n$	9	0
$WS_s > WN_s$	8	1
$WS_s > RN_n$	9	0
$WS_s > WN_n$	9	0
$WS_n > RN_n$	7	2
$WS_n > WN_n$	7	2
$WS_s > WS_n$	9	0
$WN_s > WN_n$	9	0
$RS_s > WS_n$	9	0
$WN_s > RN_n$	8	1
$RS_s > WS_s$	*7	1
	Total 109	7

*No difference in mean GSR could be detected between RS_s and WS_s for subject GW.

We have now analyzed the relative strengths of all processes underlying shock syllables for each of the 15 possible combinations of the experimental categories. Definite inequalities can be stated for 13 of these comparisons. According to the symbolic-report hypothesis, as it has been formulated in this section, the category for which the shock-syllable processes are stronger will have the greater GSR. This hypothesis can now be tested by comparing the 13 inequalities of shock-syllable strength with inequalities of GSR based upon the experimental data reported by Lazarus and McCleary. Table 2 presents this comparison. The left-hand column states the 13 inequalities derived above. The column on the right gives the number of observers in the Lazarus and McCleary study whose inequalities of GSR are in the predicted direction and the number in the opposed direction. The data are taken from their figure (p. 119). For none of the 13 inequalities do more than two of the nine observers show differences in GSR opposed to the predictions based on the statistical form of the symbolic-report hypothesis. Of the 116 inequalities that can be established from the Lazarus and McCleary data, only seven run counter to prediction.

Table 3 presents a similar tabulation for the two unreliable inequalities. These compare categories that cannot be expected to show reliable differences in GSR for the small numbers of observations that have been made. A chance distribution of GSR inequalities should therefore be found.

TABLE 3

TESTS OF THE TWO UNRELIABLE INEQUALITIES
(Arranged as in Table 2)

A	B	
	Yes	No
$RN_n \approx WN_n$	5	4
$WS_n \approx WN_n$	4	5
	Total 9	9

The data show just such a distribution: the nine inequalities of GSR for each of these pairs of categories are divided as equally as possible between those for which one category has the larger GSR and those for which the other category has the larger GSR. Of the 18 inequalities of GSR that can be set up from the experimental data, nine are in the same direction as the unreliable inequality of shock-syllable strength, and nine are in the opposed direction. Again the data are consonant with the analyses of shock-syllable strengths.

LOWER-ORDER REPORTS

Those assumptions about the recognition process that do not involve GSR can be tested directly by an experiment in which the observer reports more than one syllable following each exposure. Only the theory of these phenomena will be discussed here. The relevant data could be obtained from a modification of the recognition phase of the Lazarus and McCleary experiment in which the observer would be instructed to list, following each exposure, the m different syllables in the order in which they seemed to him most likely to have been the exposed syllable. Alternatively, the subject can be told that his initial report was incorrect and asked to give a different report. If this too is incorrect, he can be asked to give a third, etc., until he makes the correct report.

The initial or first-order report in either of these series can be identified with the single report usually recorded in a recognition experiment. We have defined the *strongest underlying syllable process* by this first-order report. Now suppose we eliminate this strongest syllable process from the alternatives available to the observer by asking him to give a different second-order report. The syllable then reported must, by the same definition, correspond to the *strongest remaining syllable process.* Thus the second-strongest syllable process will correspond to the second-order report. Similarly, if we eliminate $r - 1$ alternatives on the basis of the observer's reports, the rth-order report should correspond to the rth-strongest syllable process.

In the argument that follows we shall derive an expression for the probability that an exposed syllable will be emitted on the rth-order report if it has not been emitted on a previous report in terms of its first-order probability. We shall assume as before that a syllable process has a definite strength following each exposure and that this strength is a random variable over a series of exposures of the same duration. We shall also continue the simplifying assumption that the average strengths of all unexposed syllables differ only negligibly from one another, an assumption provided for by the counterbalanced design of the Lazarus and McCleary experiment.

Consider a matrix of the probabilities of emission for each of m possible syllables on r successive reports. Let i be the exposed syllable and α, β, γ, . . . be the $m - 1$ unexposed syllables. Table 4 shows such a matrix for $m = 6$. We wish to show that our assumptions determine

TABLE 4

MATRIX SHOWING THE rTH-ORDER PROBABILITIES OF EMISSION FOR EACH OF SIX SYLLABLES

In this example it is assumed that $P_1(i) = .5$ and that the unexposed syllables α, β, . . ., ϵ are emitted on first-order, second-order, . . ., fifth-order report, respectively.

i	r					
	1	2	3	4	5	6
	5/10	5/9	5/8	5/7	5/6	5/5
α	1/10	0	0	0	0	0
β	1/10	1/9	0	0	0	0
γ	1/10	1/9	1/8	0	0	0
δ	1/10	1/9	1/8	1/7	0	0
ϵ	1/10	1/9	1/8	1/7	1/6	0

the entire matrix when only the first-order probability of the exposed syllable i is given.

Let us consider the first-order reports. Let $P_1(i)$ be the probability that the exposed syllable is emitted as the first-order report. Then the probability that one of the $m - 1$ unexposed syllables will occur as the first-order report is $Q_1(i) = 1 - P_1(i)$. The probability that any particular unexposed syllable, say α, will be the first-order report is $Q_1(i)/(m - 1)$, since the strengths of all unexposed syllables are assumed to be equivalent. The column for $r = 1$ in Table 4 has been worked out accordingly for $P_1(i) = .5$.

Next let us consider those exposures for which the first-order report is an unexposed syllable. By our assumptions, second-order reports will be determined by the same distribution of syllable strengths that determine first-order reports, save that the strength of the syllable emitted on first-order report must be zero. Suppose α to be the unexposed syllable emitted on first-order report. The second-order probabilities of all remaining syllables will differ from their first-order probabilities only by virtue of the fact that it is impossible for α to be emitted on second-order report, while the probability that α will be emitted on first-order report, $P_1(\alpha)$, is greater than zero. Consider a large number of exposures N for which there will be $n_i = NP_1(i)$ first-order reports of the exposed syllable and $n_\alpha = NP_1(\alpha)$ first-order reports of the unexposed syllable α. We now ask, what is the probability of i if all α responses are eliminated? The answer, $n_i/(N - n_\alpha)$, gives us the probability $P_2(i)$ that the exposed syllable will be given on second-order report if α is the first-order report. The second-order probabilities of each remaining unexposed syllable can be found in the same way. For the unexposed syllable β, for instance, $P_2(\beta) = n_\beta/(N - n_\alpha)$, where $n_\beta = NP_1(\beta)$. These probabilities of second-order report are illustrated in the column for $r = 2$ in Table 4. Since all unexposed syllables are equally probable, the same solution would be reached if the syllable emitted on first-order report were some unexposed syllable other than α. We can therefore let the general symbol j_1 represent whatever unexposed syllable is emitted on first-order report. It follows that

$$P_2(i) = \frac{n_i}{N - n_{j_1}} = \frac{NP_1(i)}{N - NP_1(j_1)} = \frac{P_1(i)}{1 - P_1(j_1)}. \quad [5]$$

The reasoning upon which this equation is based can be seen more clearly when stated in terms of a conventional statistical device. Suppose there to be an urn containing N balls which differ only with respect to color. There are balls of m different colors in the urn; n_i of them are white and the remaining balls are divided equally among the $m - 1$ remaining colors α, β, γ, $\cdots$. After each draw from the urn we remove all balls hav-

ing the color of the ball that was drawn. Suppose the initial draw turns up a ball of color α. All balls of that color are thereupon removed from the urn, leaving only $N - n_\alpha$ balls in the urn for the second draw. Then the probability of obtaining a white ball on the second draw will be $n_i/(N - n_\alpha)$. Since $n_\alpha = n_\beta = \cdots$, it can be said in general that the probability $P_2(i)$ of obtaining a white ball on second draw if the first draw yields a nonwhite ball is $n_i/(N - n_{j_1})$, where $n_{j_1} = n_\alpha = n_\beta = \cdots$ is the number of balls having the same color as the ball which was first drawn from the urn. Equation 5 follows immediately from the definitions $P_1(i) = n_i/N$ and $P_1(j_1) = n_{j_1}/N$.

The same reasoning can be followed to show that

$$P_3(i) = \frac{P_2(i)}{1 - P_2(j_2)}, \qquad [6]$$

where $P_2(j_2)$ is the second-order probability of whatever unexposed syllable is emitted on second-order report. Third-order probabilities calculated on the assumption that the first- and second-order reports are the unexposed syllables α and β, respectively, appear in the column for $r = 3$ in Table 4. In general, the probability $P_r(i)$ that the exposed syllable will be emitted on the rth-order report if unexposed syllables have been emitted on all preceding reports is

$$P_r(i) = \frac{P_{r-1}(i)}{1 - P_{r-1}(j_{r-1})}. \qquad [7]$$

For reports of a given order, say the rth, the probability of any unexposed syllable, $P_r(j_r)$ will equal the complement of the probability of the exposed syllable, $Q_r(i) = 1 - P_r(i)$, divided by the number of equally-likely unexposed syllables. On the rth order of report the number of unexposed syllables whose strengths are greater than zero will be $(m - r)$, since $(r - 1)$ of the original $(m - 1)$ unexposed syllables will have been reduced to zero strength by virtue of having been emitted on an earlier report. Then for the $(r - 1)$th report

$$P_{r-1}(j_{r-1}) = \frac{Q_{r-1}(i)}{m - (r - 1)} = \frac{1 - P_{r-1}(i)}{m - (r - 1)}. \qquad [8]$$

Substituting into Equation 7, we obtain

$$P_r(i) = \frac{P_{r-1}(i)}{1 - \dfrac{1 - P_{r-1}(i)}{m - (r - 1)}} = \frac{(m - r + 1)P_{r-1}(i)}{(m - r) + P_{r-1}(i)}. \qquad [9]$$

Similarly,

$$P_{r-1}(i) = \frac{P_{r-2}(i)}{1 - P_{r-2}(j_{r-2})} = \frac{(m - r + 2)P_{r-2}(i)}{(m - r + 1) + P_{r-2}(i)}. \qquad [10]$$

To obtain $P_r(i)$ in terms of $P_{r-2}(i)$ we substitute for $P_{r-1}(i)$ in [9] its value in [10] and simplify. This gives

$$P_r(i) = \frac{(m - r + 2)P_{r-2}(i)}{(m - r) + 2P_{r-2}(i)}. \qquad [11]$$

When this reduction is carried out a times,

$$P_r(i) = \frac{(m - r + a)P_{r-a}(i)}{(m - r) + aP_{r-a}(i)}. \qquad [12]$$

Proof of Equation 12 is by mathematical induction. The $(a + 1)$th reduction is defined by the solution of Equation 9 for $P_{r-a}(i)$ in terms of $P_{(r-a)-1}(i)$. Substituting for $P_{r-a}(i)$ in Equation 12 and simplifying, we have

$$P_r(i) = \frac{[m - r + a][m - (r - a) + 1]P_{(r-a)-1}(i)}{m - (r - a) + P_{(r-a)-1}(i)} \Big/ (m - r) + \frac{a[m - (r - a) + 1]P_{(r-a)-1}(i)}{m - (r - a) + P_{(r-a)-1}(i)}$$

$$= \frac{[m - r + a][m - r + a + 1]P_{r-a-1}(i)}{[m - r][(m - r + a) + P_{r-a-1}(i)] + a(m - r + a + 1)P_{r-a-1}(i)}. \qquad [13]$$

With further simplification the denominator becomes

$$(m - r)(m - r + a) + P_{r-a-1}(i)[(m - r) + a(m - r + a + 1)] = (m - r)(m - r + a) + P_{r-a-1}(i)[(a + 1)(m - r + a)].$$

The $(m - r + a)$ terms cancel with the same term in the numerator of Equation 13, and, after bracketing,

$$P_r(i) = \frac{[m - r + (a + 1)]P_{r-(a+1)}(i)}{(m - r) + (a + 1)P_{r-(a+1)}(i)}. \quad [14]$$

The conditions for proof by mathematical induction are fulfilled, since Equation 14 is the equivalent of Equation 12 for $a = (a + 1)$.

The desired solution for $P_r(i)$ in terms of $P_1(i)$ is a special case of Equation 12. Let $a = (r - 1)$, then

$$P_r(i) = \frac{[m - r + (r - 1)]P_{r-(r-1)}(i)}{(m - r) + (r - 1)P_{r-(r-1)}(i)} = \frac{(m - 1)P_1(i)}{(m - r) + (r - 1)P_1(i)}. \quad [15]$$

Equation 15 states the probability that the exposed syllable will be reported as at least the rth-order report *in terms of its probability as the first-order report.* It can therefore be of great value in the theoretical analysis of tachistoscopic experiments. It provides a direct and independent test of the assumptions from which the Lazarus and McCleary data have been deduced here.[2] What is of greater importance, it provides a quantitative, predictive test of experimental measurements of the stochastic variables that are assumed to underlie perceptual report. Such a formulation can be of value not only in the further analysis of subception phenomena, but in the development of a general theory of the role of response processes in perception.

It will be observed that Equation 15 depends upon the assumption that the various unexposed syllables have equal first-order probabilities. Actually the equation will give fairly accurate solutions when these probabilities are not equal, unless the differences are very extreme. Such cases are not likely to arise in experiments making use of nonsense syllables. If an exact solution should be desired under these conditions, it can be obtained by substituting experimental values for $P_1(j_1)$, $P_2(j_2)$, $\cdots$ $P_r(j_r)$ and carrying out successive reductions in the manner of Equations 9 and 10. The only additional data required by this type of solution are the first-order probabilities of all unexposed syllables.

[2] Experimental proof that the probability of correct report is greater than chance appeared after this paper was completed (Bricker, P. D., & Chapanis, A. Do incorrectly perceived tachistoscopic stimuli convey some information? *Psychol. Rev.*, 1953, **60**, 181–188). The data reported there, however, are not extensive enough for a quantitative test of Equation 15.

SUMMARY AND CONCLUSION

Since the statistical assumptions put forward here account for the subception effect, it is unnecessary to postulate the existence of a subception process ("discrimination without awareness"). Although postulation of a subception process cannot be excluded as an alternative interpretation, the present assumptions can be shown to possess several advantages over it.

1. The present assumptions are consistent with what has been called the symbolic-report hypothesis. This hypothesis is a general principle of wide application to perceptual phenomena. As long as this hypothesis can be shown to hold without exception, it provides an empirical law of great analytic value.

2. From the present assumptions it is possible to deduce some 13 definite inequalities of GSR between the Lazarus and McCleary experi-

mental categories, whereas only one inequality follows rigorously from the subception hypothesis itself.

3. The present assumptions assign a definite significance to exposure time as a parameter of the inequalities of GSR. Exposure time will affect inequalities differently according to the categories being compared. Although the Lazarus and McCleary data are insufficient for this kind of analysis, the effects can be tested in further experiments.

4. The present assumptions yield, in addition to the interpretation of GSR inequalities, quantitative deductions of syllable probabilities for lower-order reports. The subception hypothesis offers no interpretation of these data.

5. The present assumptions are subject to quantitative formulation on the basis of experimental data. The two basic relationships needed are (*a*) the function relating average strength of a syllable process to the duration for which the corresponding syllable is exposed, and (*b*) the function relating mean GSR to the average strength of the processes underlying shock syllables. Since the average strength of a syllable process is defined in terms of the probability of that syllable's report, both of these functions can be determined by experiment.

All these advantages may be subsumed under the statement that the statistical form of the symbolic-report hypothesis possesses far greater potentiality for the analytic description of perceptual phenomena than does the subception hypothesis. Even if those implications which cannot yet be compared with experimental data should fail to be confirmed by later tests, this general advantage would still rest with assumptions of the type of the present ones. For only when *ad hoc* postulations like the subception hypothesis are replaced with theoretical concepts having the mathematical properties of variables can rigorous description be given to a wide range of phenomena.

REFERENCES

1. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies' "Emotionality and perceptual defense." *Psychol. Rev.*, 1950, **57**, 229–234.
2. HOWES, D. H., & SOLOMON, R. L. Visual duration threshold as a function of word probability. *J. exp. Psychol.*, 1951, **41**, 401–410.
3. LAZARUS, R. S., & MCCLEARY, R. A. Autonomic discrimination without awareness: a study of subception. *Psychol. Rev.*, 1951, **58**, 113–122.
4. MCGINNIES, E. M. Emotionality and perceptual defense. *Psychol. Rev.*, 1949, **56**, 244–251.
5. MCGINNIES, E. M., COMER, P. B., & LACEY, O. L. Visual recognition thresholds as a function of word length and word frequency. *J. exp. Psychol.*, 1952, **44**, 65–69.
6. PRESTON, K. A. The speed of word perception and its relation to reading ability. *J. gen. Psychol.*, 1935, **13**, 199–203.
7. SOLOMON, R. L., & POSTMAN, L. Frequency of usage as a determinant of recognition thresholds for words. *J. exp. Psychol.*, 1952, **43**, 195–201.

(Received April 1, 1953)

Psychological Review
Vol. 61, No. 2, 1954

REINFORCEMENT AND EXTINCTION PHENOMENA[1]

JACK L. MAATSCH

Michigan State College

This paper will be concerned with a redefinition of the concept of reinforcement and its application to a comprehensive interference theory of inhibitory phenomena. These two distinct positions are part of a general theory of behavior under development by the author. However, from the standpoint of the more conventional conceptions of learning, it is felt that these two positions may stand by themselves and serve as a partial theoretical background for a series of experimental studies (1, 11, 12, 14). These studies may be considered in the light of applications of the general theory and particularly of the positions here developed.

Abstraction and the Notion of Reinforcement

The notion of reinforcement is an abstraction. It is a construct which has a definite role in a system of abstract ideas (theory); and to be meaningful as an idea, or as a linguistic term denoting the idea entertained, the definiens must contain, or be reducible to, symbols denoting observables. If the idea is an abstraction from objects of the *thing* level, e.g., animal or wood, then the simple ostensive or denotative mode of definition would suffice. If, however, the abstraction is one pertaining to unobservables, as in the case of reinforcement, then the operational definition is the proper form of definition to secure intersubjective agreement concerning the possibility of entertaining the idea in question (13).

We may distinguish the *theoretical* properties of a construct (its functional relationship to other abstractions in a theory) from a given operational definition of that idea. The notion of reinforcement carries the connotation of cause.[2] Reinforcement causes learning. As a causal agent, then, the assertion or denial of the idea is tantamount to the prediction of learning or its absence in a given situation. Reinforcement, once asserted, must specify what is learned, that is, what response is learned to what stimuli. Therefore, we may expect that an adequate operational definition of the notion will determine the grounds for asserting or denying learning; and further, it will precisely specify what response is learned to what stimuli.

We wish to defend the theoretical idea of reinforcement; but we wish to deny contemporary operational definitions of

[1] I wish to express my gratitude for invaluable aid received through discussion with Professor M. Ray Denny, R. A. Behan, R. H. Davis, H. M. Adelman, O. Smith, R. Martindale, and M. Goer.

[2] The term "cause" is here used in the sense usually employed within philosophy of science. To assert that such and such a state of affairs *causes* learning, is to assert that there exists a more molecular chain of events (a neurophysiological chain of events) which links the observed "state of affairs" with the observed modification of behavior. Therefore, the definition of reinforcement should define the parameters necessary to produce a neurophysiological chain of events which makes reasonable the relationship between such and such a situation and the observed modification of behavior. Thus it may be argued that the notion of sheer "contiguity" specifies only one of the parameters for learning, that the notion of "drive reduction" specifies an erroneous, or at least unparsimonious, parameter, and finally, that "elicitation," the hypothesis to be proposed, does delineate an acceptable and parsimonious set of parameters.

this idea, specifically the "drive reduction" hypothesis and the "contiguity" hypothesis. The drive reduction hypothesis would seem to be empirically false (5, 6, 18, 19), neurologically implausible as the cause of learning (5, 6), and operationally inconsistent. For example, compare the definitions and applications of primary and secondary reinforcement (10). "Contiguity" as formulated by Tolman (22) and Guthrie (4) is universally applicable and therefore meaningless within the present system. One may never *deny* the occurrence of learning to contiguous stimuli; therefore strict adherence to a contiguity hypothesis yields false predictions when utilized in conjunction with the habit concept. One must predict that the animal will tend to learn to do what it did on the previous trial. Thus, a redefinition of the construct reinforcement would seem to be in order at the present time.

THE ELICITATION HYPOTHESIS

A reinforcement (E) will occur whenever there occurs a stimulus (s) or a stimulus complex (S) that elicits a characteristic response (r). Given the occurrence of a reinforcement, there will result an increment to a tendency (${}_sH_R$) for that complex to evoke a member of that response class (R).

There are several features of the definition which require some elaboration. First, the critical operation to secure or to assert a reinforcement is that of producing or causing some identifiable member of a response *class* to occur. For example, the experimenter makes an animal hungry so that it will approach and eat a pellet of food. Secondly, a member of *that* response class will reoccur with increasing probability and decreasing latency, and a member of that response class will occur if stimulus generalization occurs in a behavior sequence. The anticipatory occurrences of responses which are produced by stimulus generalization of the habit to places temporally antecedent to the eliciting stimulus complex may provide the basis for the explanation of extinction phenomena. This matter will be discussed below.

The operational determination of the reinforcing stimulus, where seemingly indeterminant, e.g., the latent learning catastrophe, may proceed by the time-honored methods of determining causal agents if one is wary of the inferential pitfalls of such methods (3, ch. 13). Within the present formulation of reinforcement, to test for latent learning in an experimental situation that excludes reinforcement, one merely eliminates the stimulus or stimulus complex that produces the response class to be "latently" learned.

One additional point concerning the definition should be made, namely, the lack of an operational distinction between primary and secondary reinforcers. Both elicit identifiable responses; hence both reinforce. One type is prior to and not dependent upon previous learning, e.g., shock, while the other *is* contingent upon previous learning, e.g., discriminated stimuli. There is, however, a difference in the effects of using these different types of reinforcing agents to produce new learning. This leads us to a third type of reinforcer which is also contingent upon past learning and upon the presentation or the occurrence of "secondary" reinforcers. We will tentatively define frustration as follows: ***Frustration stimulation*** (s_f) will occur in a learned sequence whenever the elicitation of a learned response (CR) results in the occurrence of a stimulation complex (S) interrupting performance of the learned sequence. s_f will elicit members of a characteristic class of responses (Rs_f).

The definitions of E and s_f, although structurally similar, are of a quite different nature with respect to the abstract idea entertained. Frustration is a par-

ticular kind of stimulation occurring in a specific type of situation; and learning with respect to the situation must have occurred to some extent before s_f will occur. Care should be taken to distinguish s_f from emotion, here considered an unconditioned response to other types of situations, although frustration may result in "emotional" behavior. Further, s_f has the property to *elicit* a class of characteristic behavior. Therefore, the construct may be considered an axillary of E, specifically *demanding* the assertion of E. Since frustrating situations are reinforcing, the stimulus complex which initiates s_f will be conditioned to Rs_f.

The reactions to s_f will vary as the particular physical situation varies. For example, s_f occurring after an anticipatory turn into a cul results in responses characterized by "recoil," whereas s_f occurring in an enclosed goal box results in behavior grossiy described as "emotional." Habits based upon s_f behave as any habit and interact with the original habit frustrated. It is this interaction of habits based upon two different sources of reinforcement, e.g., *US*, s_f, which seems to suggest an explanation of some complex extinction phenomena.

APPLICATIONS

Since the selective marshaling of positive experimental evidence is not particularly impressive or parsimonious, perhaps it will be better to apply the hypothesis in a few general cases.

Shock and learning. In experiments involving shock, extreme care should be taken to observe exactly what kind of responses are actually being elicited from each animal. These are the responses that will be learned. For example, consider the experiment by Brogden, Lipman, and Culler (2) which has played an important role in the development of the two types of conditioning formulated by Hilgard and Marquis (7). Under the classical procedure, animals are shocked with every presentation of a buzzer while in a revolving cage. The results of this procedure show, first, a steady but slow increase in "anticipatory running," and then, in the latter stages of the experiment, a steady decay in anticipatory running. On the other hand, during instrumental escape administration, shock is withheld if the running response appears in anticipation of administration of shock. Here the results show rapid, unhindered learning.

Consider the effect of the two methods of administration of shock upon the responses elicited (17). With the classical procedure, the running response is elicited during the early stages of the experiment, for the animal is relatively stationary. As this response begins to appear in anticipation of shock, the shock appears while the animal is in motion. The response elicited, if anything, disrupts or interferes with the previously learned running response. Thus, the classical procedure, which does not assure consistent effects of shock throughout the course of learning, results in first an initial rise in the learning of "anticipatory running" and then gradual disruption of the response to be learned. Of course, new responses are now being elicited in anticipation of shock. On the other hand, with the instrumental procedure, shock is withheld if the running response is elicited by the buzzer. Thus, the response is protected from disruption by shock. Unhindered learning does result with the instrumental procedure. From the present position, the distinction between the two kinds of learning with respect to shock would seem to be theoretically superficial.

Two-factor theories of reinforcement. The conditioning of internal and autonomically produced reflexes must be treated somewhat differently from the conditioning of skeletal responses. A

characteristic of internal reflexes, not shared by the instrumental skeletal response, is that of being *elicited* by the initial phases or *onset* of the unconditioned stimulus and of being maintained throughout the duration of shock. If shock persists for a period of time, the identical internal reflex persists, while a wide variety of skeletal responses may appear throughout the duration of shock. Now if the external stimulation complex is markedly changed *during* the continued presence of shock, the stimulus complex associated with onset of shock may acquire the property to elicit internal responses but not the ensuing instrumental escape response. These internal reflexes would, however, form a distinct and persistent internal stimulation pattern which would be part of the total stimulation complex eliciting the skeletal response instrumentally made to escape shock. Therefore, with the continued repetition of the sequence, elicitation of internal responses without the elicitation of the instrumental skeletal response would become increasingly improbable.

When skeletal responses are made to continued presence of shock a different hypothesis emerges. A response elicited by a given stimulation complex is a response reinforced to that complex. However, a second and different response elicited to the same external complex would also be reinforced, the second response replacing the first as the response most likely to occur subsequently if the shock and thus the sequence of events were terminated. In other words, the last response elicited by a given stimulation complex will be the first to occur to that complex on subsequent occasions. This principle was first applied by Guthrie (4) and will be used in a similar manner in applying the elicitation hypothesis.

Through the above analysis it is seen that "drive reduction" as presently applied within the Hullian framework has only a tenuous correlation with the response to be learned. When considering the differences in the analysis of the learning of the two types of responses, internal and skeletal, it becomes apparent that two types of learning (7) or two types of reinforcement (15, 20) could easily be posited. "Contiguity" with the onset of noxious stimuli or "contiguity" with the elicitation of *internal* responses becomes a special case of the application of the elicitation hypothesis as does the reinforcement of the terminal *skeletal* response through "drive reduction."

Classical conditioning. The learning of a behavior sequence originates with the involvement of an *S* complex with a reinforcing stimulus. The first influence of a reinforcement is upon stimuli coexisting spatially and temporally. The effect of a reinforcement is upon a habit involving that complex with the response class produced. Thus, irrelevant (conditioned) stimuli come to stand in the same relation to response as does the unconditioned stimulus. This is of course similar to the substitution hypothesis as elaborated by Hilgard and Marquis (7). Because of reinforcement of trace elements present in the eliciting complex and the similarity of internal and external stimulation, i.e., hunger pangs, gray alley and goal box, constant illumination, etc., on subsequent trials, spatially and temporally prior complexes will elicit R_g. In classical conditioning, this response class is the molecular internal unlearned salivary reflex elicited by meat powder, while the temporally antecedent complex involves a ticking metronome as a discriminative stimulus.

Natural approach and avoidance. The analysis thus far is also descriptive of the learning of natural sequences which terminate with pleasant or noxious unconditioned stimuli. On subse-

quent trials, S's antecedent to S_o (the eliciting complex) come to elicit R_o (the elicited response class). In case R_o is an approach class, S's antecedent in time to S_o elicit an approach response which results in the next S complex which will in turn elicit approach responses. This sequence of events if not interrupted will automatically result in S_o and termination of the sequence. If R_o is an avoidant class, then temporally antecedent S's will elicit responses which do not lead to the complex more approximate to S_o. This sequence of events, if unhindered, automatically leads to the elimination of the sequence.

An interference theory of inhibition. This position postulates that extinction is a result of the learning of responses which are elicited by the omission of the reinforcing stimulus in a learned situation. This new response tendency interferes with the original response tendency. Complete extinction, then, is viewed as an unstable equilibrium of two competing response tendencies to a common stimulus complex (11).

By applying Hull's stimulus-patterning corollary V (9, p. 379) to the extinction situation, the following law may be derived: any change in the stimulus situation which was present during learning will result in the learning of competing responses (Rs_f) to a discriminative stimulus complex (Ss_f). Where minimal change of the stimulus complex occurs, e.g., omission of reward only, extinction will be described as *relearning*. Where marked changes occur at the beginning of extinction, we will refer to the extinction process as *discrimination learning*. It follows from corollary V (9, p. 368) that the greater the stimulus difference between the learning and extinction situations, the faster the discrimination and, as a consequence, the faster the so-called "inhibition" of the original response due to learning of an incompatible response to frustration. The introjection of massing procedures when learning was spaced or the elimination of "secondary reinforcing" (discriminative) stimuli may serve as an example of this prediction.

In the following section, we should like to try to answer specifically the difficulties facing the formulation of this type of theory as posed in a review of theories of learning:

> 1. . . . what makes the competing response occur initially, and how does it eventually become stronger than the originally learned response? (21, p. 713.)

The competing response is initially elicited by frustration resulting from the removal of the reward in the learned sequence. This state of affairs, by definition, results in the reinforcement of "that response" to "that stimulus complex." During extinction, as long as the original response tendency is stronger, frustration will again occur and the competing response tendency will be further strengthened. An unstable equilibrium is the end result of "complete" extinction with the new interfering tendency slightly the stronger of the two tendencies.

> 2. One of the most embarrassing findings is the fact that the massing of trials has opposite effects on conditioning and extinction (21, p. 713).

Original conditioning usually concerns the development of habits to new stimuli in an unfamiliar situation, while extinction involves the learning of new responses in a previously learned situation. As a consequence, care should be taken when generalizing from the effects of variables in the one situation to effects of the variable in the other situation. Massing procedures often generate unconditioned emotional responses. These would tend to interfere with the learning of a new response. However, during extinction these same emotional

responses would facilitate the extinction of the learned response. The effect of massing procedures upon rate of extinction should also depend upon the kind of procedure employed during learning. When learning is spaced, and it usually is, massing of extinction trials results in a discriminative extinction situation, and extinction of the original response should be facilitated.

3. . . . the absence of positive correlations between conditioning and extinction measures . . . (21, p. 713).

In order to derive positive correlations between the rate of learning and the rate of extinction utilizing an interference theory it is necessary to assume: (*a*) that the two situations are identical *and they are not here considered to be*, (*b*) that there exists some general "learning ability," which does not seem to exist at all (8, p. 635), and (*c*) that this ability is the *only* variable contributing to rate of learning and extinction, which is at least questionable. However, we might assume that rate of learning is indicative of adjustment to the experimental apparatus and procedure which is in turn related to some emotionality factor. Here we have a general factor common to both learning and extinction. However, this factor could result in slow learning but fast extinction since the occurrence of emotional responses would result in additional interference to the original response. In this case the interference theory would generate the prediction of negative correlations. In general, the absence of positive correlations may be viewed as a confirmation of, and not an inconsistency with, predictions stemming from a pure interference theory.

4. . . . the differential effects of an extra stimulus on responses extinguished to different degrees . . . (21, p. 713).

The phenomenon of disinhibition seems to be a rather complex phenomenon and offers difficulty to rival theories of inhibition (20, pp. 96–102; 21, p. 715). However, within the present position, if we can assume that the extra stimulus is sufficiently pervasive to disrupt the stimulus complex actually eliciting the response class in question, several predictions can be made. First, in a relearning type of extinction situation, since the complex in question elicits both the learned and interfering habits indiscriminately, little or no disinhibition would be manifested. In the discriminative type of extinction situation, the extra stimulus would selectively disrupt the discriminative stimuli to which the interfering responses are being learned, allowing the original habit to manifest itself more rapidly following presentation of the extra stimulus. Granted the above analysis, we may expect in the relearning type of situation that any observed disinhibition would decrease as extinction progressed. In the discrimination type of situation, we might expect the opposite. Suffice it to say that a pure interference theory may be able to handle adequately the phenomenon of disinhibition; however, much more research in this area is needed before the phenomenon of disinhibition may be regarded as embarrassing to any theory of inhibition.

5. . . . the different types of extinction curves following massed and distributed conditioning . . . (21, p. 714).

In general, the original learning conditions and their relationship to the conditions obtaining during extinction are considered the primary determiners of the obtained extinction curves, and variations in the learning conditions will certainly produce noticeable effects upon extinction curves, e.g., partial reinforcement, variation of intertrial interval, etc.

6. . . . the extinction of a conditioned pupillary response in a curarized cat (21, p. 714).

It would seem plausible to attribute extinction of this reflex to physiological fatigue since the response in question

involves such a limited number of antagonistically operating muscles. The interference theory here employed is concerned with the prediction of gross skeletal responses of intact organisms in complex stimulus fields under normal learning conditions. In this case, physiological fatigue or inhibitory states arising from performance or work are considered irrelevant variables. This is not to deny that physiological fatigue is irrelevant, but to assert that its effect in the usual learning experiment is negligible. For experimental evidence for this assumption see (12). A more interesting problem might be to explain the learning of this response from the drive reduction position of reinforcement.

7. . . . and even the phenomena of spontaneous recovery . . . (21, p. 714).

Spontaneous recovery within this system is considered to be a function of the reinstatement of cues previously associated with reward. Since extinction is described as an unstable equilibrium between competing response tendencies, any partial reinstatement of cues which have been discriminatively associated with reward may be expected to result in a disequilibrium and consequently "spontaneous" recovery. The theory affords a system in which many predictions concerning spontaneous recovery may be generated, some of which contradict predictions stemming from the application of the theory proposed by Hull.

8. Still other difficulties confronting this type of interference theory may be found in Razran's thorough discussion of the problem (21, p. 714).

An analysis of Razran's (16) criticisms, which are experimentally founded and are not considered above, yielded no serious difficulties for the present position within the scope intended.

Having specifically answered a list of traditional criticisms of pure interference theories, I would like to summarize briefly a few of the more important experimental results we have obtained directly as a consequence of testing predictions stemming from the applications of the theoretical position elaborated above.

1. The massing procedure serves as a cue when discriminately associated with nonreward (11).

2. Response inhibition may be rapidly learned where no previous inhibition (I_R) was present (11).

3. Amount of effort expended during extinction of a molar response is not a significant factor in the extinction process (12).

4. Partial reinforcement produces discrimination learning, and the resulting resistance to extinction is a function of the degree of discrimination obtained prior to extinction, *OTE* (14). The tentative generic formula suggested by the theory for fixed ratio reinforcement in a Skinner-box situation is as follows: The number of responses to extinction resulting from fixed-ratio reinforcement equals the percentage of discrimination obtained under partial reinforcement prior to extinction times the reciprocal of the ratio of reinforcement employed times the number of responses to extinction of a maximal habit reinforced at a ratio of 1.0 (100%). Percentage of discrimination is defined as the ratio of responses to the food dish to the number of bar presses per reward, $D = 1 - \frac{A-1}{N-1}$.

5. Partial reinforcement per se does not yield increasing resistance to extinction where discrimination is impossible (14).

6. Resistance to extinction of a running response is a function of the type of response elicited by frustration (removal of reward). A directly opposing or incompatible response to frustration yields rapid extinction of the original response, *while a response compatible*

with the original running response yields little or no extinction of the original response (1).

7. The response to frustration will not extinguish without the introjection of some other frustrating stimulus blocking completion of that response (1).

Statements 6 and 7 when considered together yield a theoretical prediction and an experimental confirmation of the relation between fixation and extinction of a response tendency. A series of predictions drawn from the general theory concerning fixation are presently being investigated, with the hope of very simply demonstrating a "functionally autonomous" maladaptive running response in the rat.

SUMMARY

This paper has attempted to formulate a new reinforcement hypothesis and to defend a pure interference theory of extinction. We have attempted to indicate the parsimonious integrative power of the elicitation hypothesis and its role in producing inhibition in the hope that specific and detailed experimental evidence soon to be presented will be more meaningful since the facts obtained represent somewhat of a departure from hypotheses drawn from conventional behavior theory.

REFERENCES

1. ADELMAN, H. M., & MAATSCH, J. L. Resistance to extinction as a function of the type of response elicited by frustration. *J. exp. Psychol.*, in press.
2. BROGDEN, W. J., LIPMAN, E. A., & CULLER, E. The role of incentive in conditioning and extinction. *Amer. J. Psychol.*, 1938, **51**, 109–117.
3. COHEN, M. R., & NAGEL, E. *An introduction to logic and the scientific method.* New York: Harcourt, Brace, 1934.
4. GUTHRIE, E. R. *The psychology of learning.* New York: Harper, 1935.
5. HARLOW, H. F. Mice, monkeys, men, and motives. *Psychol. Rev.*, 1953, **60**, 23–32.
6. HEBB, D. O. *The organization of behavior.* New York: Wiley, 1949.
7. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning.* New York: D. Appleton-Century, 1940.
8. HOVLAND, C. I. Human learning and retention. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 613–689.
9. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
10. HULL, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.
11. MAATSCH, J. L. An exploratory study of the possible differential inhibitory effects of frustration and work inhibition. Unpublished master's thesis, Michigan State College, 1951.
12. MAATSCH, J. L., ADELMAN, H. M., & DENNY, M. R. Effort and resistance to extinction of the bar-pressing response. *J. comp. physiol. Psychol.*, 1954, **47**, 47–50.
13. MAATSCH, J. L., & BEHAN, R. A. Toward a more rigorous theoretical language. *Psychol. Rev.*, 1953, **60**, 189–196.
14. MAATSCH, J. L., DENNY, M. R., & WELLS, R. H. Resistance to extinction as a function of the number of blocks of fixed ratio reinforcement. *J. exp. Psychol.*, in press.
15. MOWRER, O. H. On the dual nature of learning—a reinterpretation of "conditioning" and "problem solving." *Harvard educ. Rev.*, 1947, **17**, 102–148.
16. RAZRAN, G. S. The nature of the extinction process. *Psychol. Rev.*, 1939, **46**, 264–297.
17. SHEFFIELD, F. D. Avoidance training and the contiguity principle. *J. comp. physiol. Psychol.*, 1948, **41**, 165–177.
18. SHEFFIELD, F. D., & ROBY, T. B. Reward value of a nonnutritive sweet taste. *J. comp. physiol. Psychol.*, 1950, **43**, 471–481.
19. SHEFFIELD, F. D., WULFF, J. J., & BACKER, R. Reward value of copulation without sex drive reduction. *J. comp. physiol. Psychol.*, 1951, **44**, 3–8.
20. SKINNER, B. F. *The behavior of organisms.* New York: Appleton-Century-Crofts, 1938.
21. SPENCE, K. W. Theoretical interpretations of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 690–729.
22. TOLMAN, E. C. *Purposive behavior in animals and men.* New York: Appleton-Century-Crofts, 1932.

(Received May 19, 1953)

Psychological Review
Vol. 61, No. 2, 1954

HYPOTHETICAL CONSTRUCTS AND INTERVENING VARIABLES [1]

ARTHUR GINSBERG

New York University

As MacCorquodale and Meehl correctly maintain, "the view which theoretical psychologists take towards intervening variables and hypothetical constructs will of course profoundly influence the direction of theoretical thought" (14, p. 95). Accordingly, it is important to examine whether the grounds for this distinction, as formulated by these authors, can be reasonably sustained, especially in face of mounting criticisms from various quarters (2, 12, 15, 16). The conclusions which the present paper will endeavor to justify are (*a*) that there are fundamental methodological and logical distinctions between what might be termed intervening variables and hypothetical constructs, but (*b*) that the grounds for these distinctions are other than those cited by MacCorquodale and Meehl. Some of the arguments presented by these authors seem to suffer from lack of clarity, and it may be because of this that the distinction has been repudiated and, even where accepted, not always with signs of adequate understanding.

Some Distinctions between Laws and Theories

1. Elsewhere (6), in somewhat closer detail, some of the characteristics distinguishing scientific laws and theories were outlined and supported against views which either denied the distinction or its importance or which maintained that laws only were the proper form which psychological knowledge should assume.

[1] The writer is grateful to Profs. Ernest Nagel and Rita W. Nealon for critically reviewing the manuscript and offering suggestions toward its improvement.

It is generally accepted that the central objective of scientific endeavor is to arrive at an understanding of the universe or various of its parts and facets. What does it mean to understand? This question is directed not upon things understood, but once removed, upon understanding itself. Its answer would afford an understanding of understanding. What are the methods, objects, and objectives of such once-removed understanding? Is it the behavioral processes undergone or methods employed by those questing understanding which is the object? Or is it the final product of such questing, the linguistic systems or logical artifacts arrived at? In either case, which behaviors or methods or which artifact outcomes are to be taken as paradigms? Connected with the foregoing, are the second-order methods involved in examining first-order methods of understanding or their outcomes identical methods, and are their outcomes of identical status or kind? That is, are the former part and parcel of a more inclusive and unitary enterprise of understanding or something quite apart and presupposed or immune? Are these second-order methods purely analytical, or are they empirical, or, perhaps, some combination? And if a combination, in what proportions and mode? What objectives are to be realized by such inquiries into the forms or tactics of knowing? Are these limited to description or may they be reconstructive in performing critical or legislative functions? And if the latter, by what warrant and to what degree?

These sundry questions suffice to illustrate the inevitability and (perhaps) perpetuity of divergent opinions concerning the meanings to be attached to and the answers to be settled upon the question "what does it mean to understand?" But it is an issue which must be joined and one which can result in enlightenment. Opinions are not all of one quality. They can be sorted out and appraised in respect to (*a*) their pragmatic significance in expediting the primary scientific objective, the understanding of things, (*b*) the degree to which they accord with the history of scientific development and the totality of common human experience, and (*c*) their measure of internal consistency and comprehensiveness.

To understand, as here to be understood, means to be able to explain or to be capable of rendering anything intelligible. Essential for any explanation is the utilization of a principle or set of principles as a major premise in an argument containing a statement or statements of fact and from which a conclusion is derived. The conclusion is explained if it can be demonstrated to be a logical consequence of the argument and if the premises are empirically certified to be true. There are two essential kinds of principles, those which are called laws and those which are called theories. The differences between them are crucial to an understanding of what it means to understand and to the distinction between intervening variables and hypothetical constructs to be advanced.

2. A scientific law is a universal, synthetic proposition whose terms, denoting abstract classes, are connected in some form of invariant association. Being synthetic means that a law refers to the universe and is thus susceptible of empirical confutation. Being universal means that a law applies without restriction to time or place: if it is true, it is true always and everywhere that that to which it pertains may exist. Accordingly, laws are extremely vulnerable, for their survival depends upon the absence of a single contrary instance. However, this vulnerability is only a surface weakness since the survival of a law through manifold tests confers a high degree of certitude upon it. Were laws in principle difficult to disprove, they could not warrant much confidence, and any law in fact so obtuse or resistant, is deservedly regarded with skepticism.

The concepts associated within a law are experimental in the sense of being explicable in terms of or directly bound up with observation sentences or protocol reports. Certification or confutation is relatively direct and unequivocal, if not in actual practice, then conceptually. But this directness and affiliation with observation is not solely a logical matter. It pertains to genetic considerations as well, since laws are commonly induced, whether by curve fitting or by a more liberal type of examination of accumulated data.

By way of illustration, consider Boyle's law, usually stated as $PV = RT$ (where R is a constant whose value depends upon the mass of the gas). Pressure, volume, and temperature are concepts which can be rigorously determined by relatively direct observational and mensurational techniques, and what they refer to is capable of being perceived. The law was adduced by Boyle as a result of manipulations upon specimens of gas whereby he noticed that with the contraction of the volume there occurred a correlated increase in the temperature and pressure in fixed ratios expressed by the law. It is a universal proposition because it applies to anything, anywhere, and at any time which happens to be a gas. That something is a gas can be readily ascertained by invoking other criteria involving laws,

e.g., the effects of passing an electrical current through it, its spectrum, its viscosity, etc. That Boyle's law is synthetic is attested to by its susceptibility to confutation; that it is true, by the absence of such confutation despite repeated exposure to a variety of tests, by its consonance with other principles, and by its practical utility.

The law signifies that alterations in at least one of the properties comprehended must result in correlate alterations in the others in predictable proportions. The actually observed instances of this invariant association are explained by showing them to comprise a subset of a more inclusive range of permissible values. And we understand the mechanical behavior of a concrete instance of gas when we are capable of providing such an explanation.

3. In contrast to a law, a theory is not a product of inductive generalization nor is it constituted by concepts which are explicable in observation terms alone. Theories are *invented*, highly abstract and general systems of concepts whose association with observations or facts is indirect rather than explicit. No amount of data inevitably determines a theory in the same way as it determines a law. Presuming the requirement of maximal mathematical simplicity, a finite set of data can be so extrapolated and interpolated as to limit the possible form or expression of the law to a few or even one case. For this reason, laws are generally highly resistant to change, a fact which is amply borne out by the history of science. It is quite otherwise with theories. Because their association with observations or facts is relatively remote, theories are notoriously mutable in time (in greater or lesser degree), and at any one time, there is likely to exist more than one theory for a given domain of subject matter. For example, in order to rectify certain deficiencies in Planck and Bohr's original quantum theory, Heisenberg, de Broglie, Schrödinger, Dirac, and recently Einstein have formulated alternate theories all of which are mathematically equivalent but differ conceptually. Such a state of affairs should be reassuring to the psychologist who is disturbed because more than one theory of learning is presently contending for general approval.

Typical theories like those of Newtonian mechanics, Einsteinian relativity, Planck's quantum, or Maxwellian electromagnetics contain ideas which are speculative in the sense of involving a constructive imaginative grasp of some root analogy, universal attribute, or system of entities of suprasensible purport or character. Such concepts apply in a variety of situations which, before the theory was invented, appeared to be independent of one another. Because of their abstractness, these concepts are differently specialized in different contexts of inquiry concerned with different phenomena. For example, in mechanics, the concept of force functions in such varied contexts as Hooke's law, Archimedes' principle, the motion or attraction of bodies in space, the law of the pendulum, laws of vibration, etc., all involving different coefficients and requiring different specializations of "force." But however manifestly different, these phenomena are revealed as instances of a common pervasive natural order by virtue of their being explained in terms of a common theory.

Distinctively theoretical concepts are postulated jointly as a system where rules of combination and transformation allow for the logical derivation of hypotheses capable of empirical test. When the boundary conditions are appropriately specified, Boyle's law can be thus derived from the kinetic theory of gases. The concepts of this theory, e.g., aggregations of molecules behaving in mechanical fashion, are not them-

selves capable of being defined in terms of observations or translated into a sense-data language. Rather, it is via such derivations as Boyle's law that the theory is empirically tested or coordinated. Owing to its logical structure, a theory may be envisaged as a matrix of concepts within which myriad channels of progression, whether lateral or hierarchical, are provided which connect diverse facts or relations within a unitary schema. Their tremendous power for organizing experiences in an intelligible and elegant manner makes theories the supreme objective of scientific inquiry, the most penetrating and satisfying form of understanding.

4. The proper grounds for distinguishing intervening variables and hypothetical constructs derive from the differences between a law and a theory. In what follows, the fundamental idea to be advanced is that intervening variables refer to concepts which are defined in terms of or comprise the ingredients of laws, whereas hypothetical constructs refer to concepts which are the constituents of a theory. The former, therefore, are capable of ostensive or extensive formulation while the latter are either primitives within a theory and thus defined intensionally, i.e., in connection with the rules of the theory, or high-level derivatives of the primitives. Considered in this light, many of MacCorquodale and Meehl's characterizations receive a more coherent and cogent grounding. But it also follows that some of their arguments cannot be sanctioned.

The criticisms of MacCorquodale and Meehl's views that will be developed will function as a means for amplifying these contentions. Because of the obscurities present in their article, MacCorquodale and Meehl are open to a number of alternative interpretations, some of which are untenable.

Do Intervening Variables Denote?

There are a number of arguments in MacCorquodale and Meehl's article (14) which indicate that as they conceive the distinction, intervening variables do not denote whereas hypothetical constructs do. The following quotations prove this supposition: hypothetical constructs "involve the hypothesization of an *entity, process,* or *event*" whereas intervening variables are simply "operational" (pp. 95–96); intervening variables "are merely names attached to certain convenient groupings of terms in [an] empirically fitted equation . . . [and] 'exist' [only] in the trivial sense that the law holds" (p. 99); our notion of intervening variables "involves nothing which is not in the empirical laws that support them" (p. 100); intervening variables are strictly mathematical except "where the verbal accompaniment of a concept . . . makes it a hypothetical construct" (p. 101); an intervening variable is "simply a quantity obtained by a specified manipulation of the values of empirical variables" (p. 103); in contrast to an intervening variable, "it is the business of a hypothetical construct to be 'true'" (p. 104); an intervening variable "is merely a shorthand summarization [while] for hypothetical constructs, there is a surplus meaning that is existential" (p. 106).

Is this a defensible way of distinguishing between classes of concepts employed in science? Are there legitimate and useful terms in science which refer to nothing or which denote empty classes?

Clearly, by definition, logical terms are nonfactual in this sense and so far as they are found in laws or theories, they fulfill a purely syntactical function. Among these are such terms as are determined by or name the choice of measuring units; formation signs such as universal or existential opera-

tors, e.g., all, some, every, etc.; connectives, e.g., and, or, implies, etc.; and punctuation marks. All these signs are linguistic conventions applied in accordance with logical-syntactical employment rules and refer to nothing which is or might be existential. It is clear that intervening variables are not intended by MacCorquodale and Meehl to represent such nondenotative terms.

Such terms as "unicorn" or descriptions like "the present king of France" lack denotative meaning since there is no class of actual existences to which they would correctly apply. However, it is not inconceivable that there should have been or might yet be a time when these terms would correctly apply to a class of existences. At present, while assertions involving these terms are either false or without practical scientific consequences, they are not logical contradictions. Accordingly, we must allow that were certain conditions to obtain, e.g., if at time t and place p one could observe an animal shaped like a horse, with a long horn in its head, fleet of foot, etc., such terms might correctly apply. We deny them denotative meaning because such meaning attribution would flout attested facts or certified laws or theories.

It is clear that MacCorquodale and Meehl do not suppose intervening variables to be analogous to or instances of fictive terms. For example, ${}_sH_R$ is interpreted as an intervening variable. Yet, were it fictive in the sense of denoting nothing or in naming an entity, process, or property whose existence could not be truthfully affirmed in any case, we would have to allow that, so far as we understood the conditions whereby ${}_sH_R$ *would* apply correctly, were these conditions actually realized, the term would acquire denotative meaning. In a certain sense, this possibility is provided for. MacCorquodale and Meehl contend that when such terms as ${}_sH_R$ are provided with a physiological interpretation, they are converted into hypothetical constructs and are therefore either true or false, i.e., they acquire denotative meaning.

It does not seem likely that MacCorquodale and Meehl would maintain that no psychological concepts denote or that, unless they were somehow connected with physiological references, psychological terms would refer to nothing existential. Yet, at least on the issue of denotative meaning, it is difficult to see why ${}_sH_R$ should be different from terms like "perception" or "memory" and the like.

The suggestion that the distinction between intervening variables and hypothetical constructs be based on the latter's representing or connoting neural processes while the former do not, has been made by Hilgard (8, p. 265) and by Tolman, at least in his latest writings (17). As thus construed, the distinction is not only a logical one, but is also based on factual considerations. From the standpoint of the distinction as adumbrated here, that the differences between intervening variables and hypothetical constructs stem from the differences between laws and theories, whether or not an intervening variable or a hypothetical construct has physiological bearings is entirely irrelevant. There can be physiological intervening variables and physiological hypothetical constructs just as well as psychological ones so far as there exist physiological laws and theories and psychological ones.

However, and more to the point at issue, whether any or all psychological concepts could be reduced to or somehow coordinated with a set of physiological concepts or with a physiological theory would not signal the acquisition of denotative meaning by terms which previously lacked it. The reverse would be closer to the truth, i.e., that a term

can be interpreted within an independently developed system would reinforce our warrant in believing the term to be factually significant.

The difference between (*a*) psychological concepts which might be or have been reduced to or coordinated with a set of physiological concepts or absorbed within a more inclusive unitary discipline and (*b*) psychological concepts which have not been or may be resistant to such convergence is not that the former have and the latter do not have a surplus meaning which is existential. To begin with, the possibility of such convergence is not a function of psychological concepts alone but also of the state of physiological knowledge or of scientific knowledge in general. Next, even where the connotation of a psychological concept is not sufficiently developed to possess physiological purport, it does not follow that the concept lacks any reference whatsoever. Thus, ${}_sH_R$ may be supposed to lack physiological purport and still denote a dispositional property or state of organisms.

This latter confusion appears to rest upon the nominalistic notion that only concrete properties or particulars are real or ostensively significant. However, abstract terms are not nonreferential. Such terms as "blackness," "docility," etc., and all concrete terms capable of having a suffix like "ness" or "ity" attached to them or of functioning in similar ways in discourse, name what some other term connotes. For example, as Lewis points out, " 'roundness' names that character or property which is essential in order that the term 'round thing' should apply. . . . For every concrete term, 'C,' there is a cognate term which denotes the significate of 'C' " (the essential property in virtue of which C is 'C') (13, pp. 41–42). These considerations also apply to abstractions which denote kinds of relations, such as belongingness or affinity, and to those which represent interpreted mathematical equations, and to dispositional properties or states.

There is a final class of terms which is often considered to be analytic, namely, terms comprising a proposition which, in a given inquiry context, functions as a nominal definition or terms which function as synonyms for other terms. Since such terms assert nothing, they contribute nothing new or significant to scientific discourse save toward achieving clarity and economy of expression. According to Hempel, "a nominal definition singles out a certain *concept*, i.e., a nonlinguistic entity such as a property, a class, a relation, a function, or the like, and, for convenient reference, lays down a special name for it" (7, p. 4).

Nominal definitions are always conventional since they are rules for the use of language involving a stipulation that a *definiendum* be synonymous with a *definiens*. However, once the *definiendum* is so defined that instances of it can be readily identified, it may function in other contexts as an explanatory or factual concept. Thus if we suppose the mathematical statement of Hull's postulate 4 (10, p. 178) to be a nominal definition such that ${}_sH_R$ functions as a synonym for $M\ (1 - e^{-k\omega})\ e^{-j'}\ e^{-\omega t'}\ (1 - e^{-i\Delta t})$, it is not the case that ${}_sH_R$ also functions analytically in the other postulates of the system. On the contrary, all other appearances of ${}_sH_R$ are factual, and so when it occurs as a variable in the formula for excitatory potential (${}_sE_R$), it must possess denotative meaning. In a word, which formula in the system is taken as defining the term may be arbitrary, but once so defined, the other instances are strictly factual. Except for the classes of logical terms noted, no other significant term in an empirical theory or system of laws *always* functions ana-

lytically or without denotative meaning; i.e, no term for which independent criteria of application exist is analytic in every context whatsoever.

A term may be said to have meaning in a number of senses or modes. (*a*) A term *denotes ostensively* when it refers to a determinate particular and *extensively* when it refers to a determinate class where assertions about the particular or class are true or where these exist. (*b*) A term *connotes* in four ways: (*i*) the *conventional* connotation of a term consists in stipulations concerning its proper usage in discourse, i.e., that meaning which is generally accepted as fixing reciprocity of understanding or ascertaining whether anything is an instance of the term; (*ii*) the *intensional* connotation of a term consists in the range of terms or propositions implied by that term; (*iii*) the *objective* connotation of a term is comprised of the range of properties or characteristics common to the universe of entities the term denotes; (*iv*) the *subjective* connotation of a term consists in all the other terms associated with the term by the user or the criterion he might entertain respecting its usage.[2]

Viewed from the standpoint of this analysis of meaning, it becomes obvious that the conventional connotation of ${}_{S}H_{R}$ does not exhaust its full meaning, and that even while the term may function analytically in one context of inquiry, to be empirically significant it must be capable of functioning factually in other contexts. And at least in these other contexts, the term will not function as a synonym but rather factually, i.e., with denotative meaning.

If logico-mathematical signs and fictive terms exhaust the range of nonfactual or nondenotative terms, and if intervening variables are neither of these, they are either nonsense or else they *do* possess denotative significance. Were they nonsense, it is inconceivable that they could perform any role in scientific inquiry. The fact is that the attribution of a number implies a measurable parameter and only things, properties, and relations can be counted or measured. While some things which do not or cannot exist can be conceived as measurable or countable, we have already excluded fictions as the references of intervening variables. Accordingly, since MacCorquodale and Meehl admit that intervening variables can be measured, it follows that they do possess denotative meaning. And this conclusion is in line with the interpretation of intervening variables alluded to, namely, that they are terms whose definition is given by the laws in which they occur, which they imply, or which they presuppose.[3]

ARE INTERVENING VARIABLES DISPOSITION TERMS?

MacCorquodale and Meehl cite the concept of resistance as a physical instance of intervening variables thereby identifying them with dispositional concepts as characterized by Carnap (**4**). However, even while MacCorquodale and Meehl profess to mean exactly what Carnap means by dispositional concept,

[2] While roughly stated, this outline is sufficiently precise for our purposes here. The four senses of connotation are not unrelated. In many analyses of meaning, the conventional and intensional senses of connotation are combined under one rubric. Again, the subjective connotation may overlap with any of the others. It may be noted that a term's conventional connotation may be rendered by an operational definition. Clearly, these do not exhaust the meaning of a term (see **6**).

[3] Bergmann (**2**) has also maintained, on somewhat different grounds, that intervening variables and hypothetical constructs cannot properly be distinguished on the issue of denotative meaning. However, he concludes that owing to this, they cannot be distinguished on any grounds whatsoever.

they depart from his analysis in crucial ways.

"Resistance," they argue, "is 'operational' in a very direct and primitive sense" (14, p. 96). In contrast to hypothetical constructs, the entire meaning of intervening variables is exhausted by the sentences about them or is reducible to sentences about impressions. The impression conveyed is that a set of reduction sentences wholly determines the meaning of an intervening variable once and for all. However, if intervening variables are dispositional concepts, this is not the case.

Carnap explicitly states that "a set of reduction pairs is a partial determination of meaning only and can therefore not be replaced by a definition. Only if we reach, by adding more and more reduction pairs, a stage in which all cases are determined, may we go over to the form of a definition" (4, pp. 450–451). And it is doubtful whether such a stage can ever be actually realized at least with all such terms. That is, particularly in the case of the more fertile dispositional terms in science, there is what Hempel calls an "openness of meaning" (7, p. 29) or an indeterminacy of application in respect to not-as-yet discovered possibilities.

In contrast to MacCorquodale and Meehl's supposition that intervening variables lack existential reference, Carnap conceived of dispositional concepts as a kind of predicate which, in the "thing-language," would mean that dispositional concepts refer to real or factual properties. According to Carnap and others who have analyzed the logical status of dispositional concepts (3, 5, 9), terms like resistance and ${}_{S}H_{R}$ refer to existences in a sense not radically different from such terms as "table" and "electron."

By disposition we mean to refer to such properties in the world as become differently manifest, in predictable ways, depending upon the complex of causal factors to which they are subjected. Corresponding to alterations in the causally effective context, the manifest aspect of the disposition may vary, whether in quality or quantity, within a determinate range of possibility. It is this range which comprises the potency of an entity so distinguished. In virtue of the potency of things, they both change and endure in intelligible fashion. But in few if any cases can we rationally predict all the causal factors which might conceivably influence the manifest aspect or appearance character of a dispositional property. There is little doubt that ${}_{S}H_{R}$ denotes a dispositional property. So far as it is a property of organisms to which we refer, the question of existence is not simply confined to whether the equation for acquisition holds. Whether any organism is capable of possessing any kind or degree of ${}_{S}H_{R}$ is a factual issue and thus one which is settled in terms of appropriate empirical evidence. Therefore, the question of whether ${}_{S}H_{R}$ exists or not is no more trivial than any other such question posed in respect to any other entity or property which can be named. The introduction of an intervening variable is thus never merely a matter of convenience even though our convenience is served by such introduction.

ARE INTERVENING VARIABLES SIMPLY CONVENTIONAL?

In their summary, MacCorquodale and Meehl conclude that "the only rule for proper intervening variables is that of convenience, since they have no factual content surplus to the empirical functions they serve to summarize" (14, p. 107). Elsewhere they maintain that apparently legitimate sets of intervening variables could be obtained by "various arbitrary groupings and combinations" of the relevant equations (p. 98).

It is generally held that no concept in science is to be regarded as factually significant or as denoting a state of affairs in the world unless it has systematic purport, unless, that is, it is involved in other principles besides that for which it was invented or introduced. *Ad hoc* concepts either undergo such enhancement of purport or else they are counted spurious or empty and are discarded. Because such enrichment in the meaning of any concept must always be expected, any conceptual analysis must provide for this possibility in allowing that no concept can receive a final and full explication on the basis of its present functions in inquiry or the existing evidence by which it is supported. This openness of meaning implies that there is always an indeterminate range of factual content implicit in any significant concept.

A brief account of some scientific history will illustrate this point. At the beginning of this century, Planck, concerned to explain certain discrepancies between classical theory and experimentally observed facts of incandescence and radiation, postulated a new universal constant of nature which he called h. The principles containing this constant imply that emission and absorption processes both occur discontinuously, being in the nature of jumps of finite magnitude. Although heralding a great revolution in scientific thought, a prominent reason why Planck's theory initially met with a skeptical reception and could not rightly be accepted with studied confidence was the limited domain of applicability of h, or its poverty of factual bearings.

However, a few years later Einstein formulated a law for the kinetic energy of an ejected electron in the photoelectric phenomenon which required h as a constant of proportionality. Since then, h has been employed in a variety of contexts and now not only furnishes the basis for explaining the intensity of radiation and the wave length for which it represents a maximum (and for which it was originally formulated), but also for interpreting the quantitative relations existing in many other cases, namely, to cite a few, the specific heat of solids, the photochemical effects of light, the orbits of electrons in the atom, the wave lengths of the lines of the spectrum, the frequency of the X rays produced by the impact of electrons of given velocities, the velocity with which gas molecules can rotate, the distances between the particles which make up a crystal, and others.

In short, concepts which are formulated to explain a given limited subject matter may prove applicable in other spheres and it is this which certifies their significance or systematic utility (cf. **7**, pp. 52–54). In the case of Planck's quantum theory, the relevance of h in a variety of situations means, in the words of D'Abro, that the theory "cannot be regarded as a mere makeshift devised for the sole purpose of interpreting the law of equilibrium radiation. We must recognize therefore that the quantum theory has uncovered a new world of physical occurrences, a world formerly unsuspected" (**1**, p. 471).

Now it might be a permissible convention to hold that so far as a concept applies only to the area for which it was formulated it is an intervening variable, but as it becomes relevant to other, previously independent, subject matters, it acquires the status of a hypothetical construct. However, it is not clear what value would come of this convention, especially since it would present problems such as fixing the point at which the conversion may be said to have occurred.

In similar vein, it often happens that a concept defined in terms of a certain law is rendered more precise and general

in being amplified by a higher-level law or theory. The concept of resistance, conceived as an intervening variable by MacCorquodale and Meehl, is a case in point. Defined in terms of Ohm's law, resistance is a constant whose value is the ratio of voltage and current, a linear function. The term denotes that property of conductors which determines the amount of current which will flow when a given amount of voltage is impressed. However, it turns out that Ohm's law is inadequate in the case of such conductors as gas tubes, where, if the voltage reaches beyond or below a certain range, resistance no longer remains a constant, i.e., the slope of the curve gradually flattens. Here resistance becomes a function of other factors not provided for in Ohm's law, e.g., the field effects of electron clouds.

It is not the case, then, that Ohm's law exhausts the meaning of "resistance" nor is it the case that in being incorporated within higher-level laws, the term acquires denotative properties which it previously lacked entirely. It must always be realized that with new discoveries or new theoretical formulations, the meaning of any concept may be revised, refined, or broadened in being incorporated within new relationships. And it is precisely this application to and corroboration by diverse, independent domains that reinforces our confidence in a concept's factual significance.

The concept ${}_{S}H_{R}$ appears to fulfill these requirements since it has been applied to such prima facie diverse areas as the learning of nonsense syllables, mazes, etc., and the development of personality and neurotic symptoms, and to all the high-level mental functions of humans. In fact, it is in virtue of the capacity of ${}_{S}H_{R}$ to be differently specialized in different areas of application that may make Hull's system qualify as a theory in the sense outlined.[4] Accordingly, postulate 4 does not exhaust the meaning of ${}_{S}H_{R}$, and, if anything, the term denotes a dispositional property of organisms and this is something real or true about certain organisms and, in turn, the world.

If the foregoing arguments are correct, it follows that so far as intervening variables are to be counted as factually significant, they are not merely names for arbitrary combinations of empirical equations. In the first place, no empirical equation can be manipulated in every logically possible way and be expected to yield statements, all of which are empirically significant or scientifically useful. Many legitimate logical or mathematical operations are proscribed because the results would be absurd or completely useless. For example, in multiple-factor analysis, it is logically possible to rotate the axis in an infinite number of ways, yet, aside from the logical requirement that uniqueness and simple structure be attained, there is the empirical requirement that the rotation accord with theoretical or factual considerations. Again, the curves by which a functional law might be represented graphically are rarely (if ever) extended in all possible directions nor are all the branches exploited. The reason is that such extrapolations would either have no empirical significance or else would contravene fact or other principles.

[4] Whether Hull's system actually is a theory in the sense outlined is problematic. First, many of the formulae are derived from curve fitting, and second, it is far from being the kind of rigorous deductive system as are the physical theories cited. The fact that ${}_{S}H_{R}$ can be defined in experimental terms (as in the formula for acquisition) is not atypical or contrary to the nature of theories. The essential point is whether it can be fully or exhaustively defined in this way. If so, Hull's system would likely qualify only as a system of laws.

Furthermore, in constructing theories, economy of postulation is the most reliable insurance against inconsistency. We mean by theoretical simplicity that the number of independent factors, properties, principles, or variables be held to a minimum compatible with the scope or level of penetration desired of the theory. A new concept may be introduced into a system only if it does not produce redundant conclusions or result in inconsistencies, and only where its inclusion will enhance the predictive power of the theory. Whether MacCorquodale and Meehl's suggested concept of cumulative reinforcement would fulfill these requirements and the others noted previously is not demonstrated by them (although it appears to coincide with recent restatements of his position by Hull [11]). At any rate, if this concept is to be included within Hullian theory, it would not be an arbitrary matter or simply one of convenience. And if included, it would denote something about organisms, something which exists, rather than being simply a name for an equation and thus referring to nothing but itself. In brief, there are always methodological and empirical considerations which proscribe attaching factual significance to all possible transformations of equations or groupings of variables. Logical validity is not identical with factual truth or significance.

One last historical point is in order. The big difficulty in establishing the conception of blood circulation was the absence of any visible connections between the terminal arteries and the veins. Harvey postulated the existence of capillaries, but it was only after the microscope was considerably improved that these were actually observed as real. In similar manner, to explain his experimental results, Mendel postulated hereditary units which only recently have been microscopically and chemically identified as genes or real entities. Even while the advent of new confirmatory evidence increased the probability of these hypotheses, it did not result in the conversion of something lacking "a surplus meaning which is existential" into something possessing such "meaning." The relation between evidence and a hypothesis is logical and not predicative.

VARIABLES

If our criticisms are correct and justified, much of the rationale behind MacCorquodale and Meehl's distinction between intervening variables and hypothetical constructs is untenable. While upholding the distinction itself, a sounder rationale is suggested by the logical differences between empirical laws and theories. An intervening variable would refer to any concept whose definition is provided by laws or a set of laws whereas a hypothetical construct would refer to any concept whose definition is provided within a theoretical system. Intervening variables or law-like concepts would thus be distinguished in being introduced by experimental or operational procedures whereas hypothetical constructs or theory-like concepts would be distinguished in being introduced by postulational procedures, e.g., by interpreting the primitive apparatus of a formal calculus. Because they are more intimately associated with observation terms, intervening variables are introduced either via explicit or conditional definitions based upon invariable associations, functional relations, or causal involvements obtaining among entities or processes, among these and their properties, or among the properties themselves. Explicit definitions are usually of the form $X = \text{def. } Y$ often considered nominal or analytic. Conditional definitions involving reduction chains are necessary for introducing dispositional terms and

take the form if X is at p, then X is q if and only if X is r (see 4 and 7 for details).

The differences in logical status between intervening variables and hypothetical constructs are therefore based upon the differences between laws and theories. Accordingly, whatever characteristics distinguish laws and theories attach to intervening variables and hypothetical constructs as well. Many of MacCorquodale and Meehl's remarks fit in with this rationale.

Understood in this way, the relations between intervening variables or hypothetical constructs on the one hand, and independent or dependent variables on the other, become a function of problematic context. Rather than describing an absolute logical or meaning property of the terms so characterized, the distinction reduces to a matter of how equations are expressed or what facet of the inclusive psychological state of affairs is being inquired into or projected as problematic. To understand how a sensory complex issues in a response may require reference to intervening processes and any law adumbrated for this purpose may consist in three differentiable components, an independent variable, an intervening variable, and a dependent variable. But what functions as an independent variable in one context may function as an intervening variable or a dependent variable in another. If it is the sensory complex as such that we are concerned with, then the stimulus object may constitute the independent variable and the sensory complex the dependent variable. What might, under some circumstances, constitute an intervening variable, e.g., a neural process, may be an independent or dependent variable in, e.g., a physiologically oriented inquiry.

In general, intervening variables (whether these be law-like or theory-like) are introduced for either of two reasons: (*a*) without such an introduction, explanation may be insuperably partial or prediction inexact or (*b*) without such an introduction, explanation may be extremely cumbersome and require a host of disjoined *ad hoc* adjustments to be kept up to date. Thus, to achieve systematic elegance or a deeper and more precise understanding of subject matter, the introduction of intervening variables or hypothetical constructs may be essential. The test for their appropriateness or correctness has been stated quite often. In essence it consists in (*a*) the possibility, not otherwise realizable, of generating new and ever more precise conclusions out of a set of laws or a theory and experimentally corroborating these and (*b*) demonstrating a logical coherence among the principles or concepts of the system. It is these requirements which Hull has often described as the secure anchoring of intervening variables on both sides to observable and measurable conditions or events (10, p. 22). This "anchoring" is both a logical and an empirical or evidential affair. Incidentally, Hull's choice of the "anchor" metaphor is most apt: if the anchor be taken as observation and mensuration sentences, and the ship as the theory, it follows that the two are not identical or interchangeable and that there are different degrees of "security."

One last point needs to be made. In contrast to the position maintained here, certain writers, e.g., Kendler (12) and Marx (15), have either repudiated the distinction between intervening variables and hypothetical constructs or deprived it of logical significance. Like other orthodox operationists, they have offered a monolithic point of view in its stead. According to them, only intervening variables are to be encouraged or permitted into the corpus of mature

science since hypothetical constructs are regarded either as metaphysical hypostatizations and thus "meaningless" or as transitional crutches at best. In fixing upon certain of MacCorquodale and Meehl's arguments concerning the supposed purely conventional, denotative-less status of intervening variables, Kendler and Marx have ignored or misread the valid arguments raised in support of hypothetical constructs as denoting. One reason for this misreading may stem from the lack of clarity in MacCorquodale and Meehl's paper, but the main reason must be attributed to a severe nominalistic bias. At any rate, the whole tenor of this paper and the one previously referred to (6) renders a more detailed examination of Kendler's and Marx's views unnecessary.

SUMMARY

Two basic questions were raised: (*a*) Are there significant logical differences between intervening variables and hypothetical constructs? (*b*) Are these correctly conceived by MacCorquodale and Meehl? The reply to the first question was made by pointing out that the differences relate to those distinguishing empirical laws from theories. The reply to the second question was made by pointing out certain ambiguities and misconceptions in MacCorquodale and Meehl's thesis, the most prominent of which was the assumption that intervening variables were simply conventional and without denotative meaning. During the course of this critical analysis, the senses or modes of meaning, the nature of disposition concepts, the function and character of abstract terms, and the kinds and functions of variables were touched upon.

REFERENCES

1. D'ABRO, A. *The decline of mechanism.* New York: Van Nostrand, 1939.
2. BERGMANN, G. Theoretical psychology. *Annu. Rev. Psychol.*, 1953, **4**, 435–458.
3. BROAD, C. D. The "nature" of a continuant. In H. Feigl & W. Sellers (Eds.), *Readings in philosophical analysis.* New York: Appleton-Century-Crofts, 1949. Pp. 472–481.
4. CARNAP, R. Testability and meaning. *Phil. Sci.*, 1936, **3**, 420–471; 1937, **4**, 1–40.
5. CHISHOLM, R. M. The contrary-to-fact conditional. In H. Feigl & W. Sellars (Eds.), *Readings in philosophical analysis.* New York: Appleton-Century-Crofts, 1949. Pp. 482–497.
6. GINSBERG, A. Operational definitions and theories. *J. gen. Psychol.*, in press.
7. HEMPEL, C. G. Fundamentals of concept formation in empirical science. *Int. Encyc. Unif. Sci.*, II, 7.
8. HILGARD, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.
9. HOFSTADTER, A. Professor Ryle's category-mistake. *J. Phil.*, 1951, **48**, 257–270.
10. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
11. HULL, C. L. *Essentials of behavior.* New Haven: Yale Univer. Press, 1951.
12. KENDLER, H. H. "What is learned?"—A theoretical blind alley. *Psychol. Rev.*, 1952, **59**, 269–277.
13. LEWIS, C. I. *An analysis of knowledge and valuation.* La Salle, Ill.: Open Court, 1946.
14. MACCORQUODALE, K., & MEEHL, P. E. On a distinction between hypothetical constructs and intervening variables. *Psychol. Rev.*, 1948, **55**, 95–107.
15. MARX, M. H. Intervening variable or hypothetical construct? *Psychol. Rev.*, 1951, **58**, 235–247.
16. SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, **57**, 193–216.
17. TOLMAN, E. C. A psychological model. In T. Parsons & E. A. Shils (Eds.), *Toward a general theory of action.* Cambridge: Harvard Univer. Press, 1952. Pp. 279–361.

(Received May 18, 1953)

Psychological Review
Vol. 61, No. 2, 1954

SOME VIEWS ON MATHEMATICAL MODELS AND MEASUREMENT THEORY[1]

C. H. COOMBS, H. RAIFFA,[2] AND R. M. THRALL

University of Michigan

We shall undertake first to review the role of mathematical models in a science and then briefly discuss the models used in classical measurement theory. This will be followed by a generalization of measurement models. Illustrations will be introduced when needed to clarify the concepts discussed.

The Role of Mathematical Models

We shall use the terms *physical objects*, *real world*, and *object system* synonymously to signify that which the empirical scientist seeks to study, including such objects as opinions or psychological reactions. The scope and content of a domain is selected by the scientist with the intent of discovering laws which govern it, of making predictions about it, or of controlling or at least influencing it.

There are potentially at least as many ways of dividing up the world into object systems as there are scientists to undertake the task. Just as there is a potential variety of object systems, so also is there a potential variety of mathematical systems. Let us describe the nature of a mathematical system:[3] A mathematical system consists of a set of assertions from which consequences are derived by mathematical (logical) argument. The assertions are referred to as the axioms or postulates of a mathematical system. They always contain one or more primitive terms which are undefined and have no meaning *in the mathematical system*. The axioms of the mathematical system will usually consist of statements about the existence of a set of elements, relations on the elements, properties of the relations, operations on the elements, and the properties of the operations. Particular mathematical systems differ in the particular postulates which form their bases. It is evident then that the variety of mathematical systems is limited only by the ability of man to construct them.

Our view of the role that mathematical models play in a science is illustrated in Fig. 1. With some segment of the real world as his starting point, the scientist, by means of a process we shall call abstraction (A), maps his object system into one of the mathematical systems or models. By mathematical argument (M) certain mathematical conclusions are arrived at as necessary (logical) consequences of the postulates of the system. The mathematical conclusions are then

[1] This paper is a chapter in a forthcoming book, *Decision Processes*, edited by C. H. Coombs, R. L. Davis, and R. M. Thrall, published by John Wiley and Sons, Inc. It is an outgrowth of a number of sessions on measurement theory in an interdisciplinary seminar on the application of mathematics to the social sciences held during the academic year 1951–1952 by Drs. Coombs, Raiffa, and Thrall. While the influences of the separate authors are sometimes distinguishable, the paper is a truly joint product. The authors are listed alphabetically. This research was carried out under Office of Naval Research Contract Nonr 374 (00) NR 041-011 and a Ford Foundation Behavioral Studies grant.

[2] Dr. Raiffa is now at Columbia University.

[3] For a more detailed discussion of the nature of mathematical systems, see (8, 12, and 13).

converted into physical conclusions by a process we shall call interpretation (I).

Let us start with a specific real-world situation $(RW)_1$ and by process A map it into a mathematical system $(MS)_1$. We can look at $(MS)_1$ as a *model* of $(RW)_1$. Looked at in reverse, we can start with consideration of $(MS)_1$ and then $(RW)_1$ can be viewed as a *model* of $(MS)_1$, and the process of going from $(MS)_1$ to $(RW)_1$ we call "realization." Thus, "realization" is the converse of "abstraction." Now, given $(MS)_1$ we might be able to find a real-world situation $(RW)_2$ such that by assigning meanings to the undefined terms of the mathematical system, the assertions about "sets of elements," "relations," and "operations" in $(MS)_1$ become identified with objects or concepts about $(RW)_2$. That is, $(RW)_2$ may be another model of $(MS)_1$, and the process of going from $(RW)_1$ to $(MS)_1$ to $(RW)_2$ often indicates subtle analogies between systems such as $(RW)_1$ and $(RW)_2$. To the mathematician who often starts with an abstract system the model is a concrete analogue of the abstract system. To the social scientist who starts with phenomena in the real world the model is the analogue in the abstract system.

In establishing a model for a given object system one of the most difficult tasks is to attempt a division of the phenomenon into two parts; namely that part which we abstract (A) into the basic assumptions or axioms of the abstract system, and that part which we relegate to the physical conclusions and which we reserve as a check against the interpretations from the abstract system. In a given object system there is no unique partition of the phenomena, and which partition is made depends on the creative imagination of the model builder. Indeed, there are models in the physical and biological sciences for which there are no experimentally verified or verifiable correlates in the real world for the undefined terms, relations, and operations in the abstract model. A similar situation prevails on the side of the abstract system; i.e., it is often possible in a given abstract system to interchange the roles of certain axioms and theorems. Thus, in a given system there is no unique method of splitting the mathematical propositions into axioms and theorems. In going from the abstract system to the object system we have the parallel processes of realization and interpretation. It is quite common to consider these synonymous; however, we prefer in this discussion to reserve the word "interpretation" for the process which maps the mathematical conclusions (rather than the axioms) into the object system.

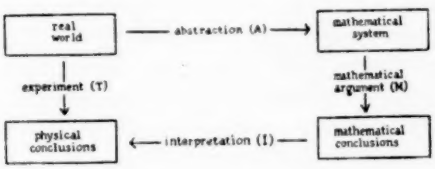

FIG. 1. The symmetrical roles of experiment and mathematics

Let us summarize briefly up to this point. Beginning with a segment of the real world, the scientist, by an entirely theoretical route, has arrived at certain conclusions about the real world. His first step is a process of abstraction from the real world, then a process of logical argument to an abstract conclusion, then a return to the real world by a process of interpretations yielding conclusions with physical meaning. But there is an alternative route to physical conclusions and this is by way of working with the object system itself. Thus, the scientist may begin with the real-

world segment in which he is interested and proceed directly to physical conclusions by a process of observation or experiment (T).

The path (T) (experimentation) from the real world to the physical conclusions needs further scrutiny. Usually in theory construction the scientist embarks on model building after he has many facts at his disposal. These facts he partitions into two parts—one part serves as a springboard for the abstraction process (A); the other part serves as a check on the model by making comparisons with these initial facts and the interpretations (I) stemming from the model. If a specific interpretation is not at variance with a fact in the initial reservoir, but at the same time not corroborated by our a priori notions of the object system, then the model *perhaps* has contributed to our knowledge of the object system. The scientist next tests this tentative conclusion by setting up a plan of experimental verification, if this is possible. Often direct verification may not be possible, and corroboration stems from examination of experimental evidence which supports claims of the model quite indirectly. That is, motivated by interpretations of the model, the scientist sets up an experimental design, obtains observations by experimentation, makes a statistical interpretation of these observations into physical conclusions, and compares the conclusions with those of the abstract route in order to appraise the model. As suggested by Frederick Mosteller[4] of Harvard University, it would be appropriate to generalize Fig. 1 as shown in Fig. 2. The route A_2EI_2 in Fig. 2 is summarized by the route T in Fig. 1. If the physical conclusions of the process A_1MI_1 are at variance with the a priori facts or with conclusions arrived at via A_2EI_2 (and if more confidence is placed in the experimental route than in the theoretical route), then the suitability of the model is suspect.

The task of a science looked at in this way may be seen to be the task of trying to arrive at the same conclusions about the real world by two different routes: one is by experiment and the other by logical argument; these correspond, respectively, to the

[4] Personal communication.

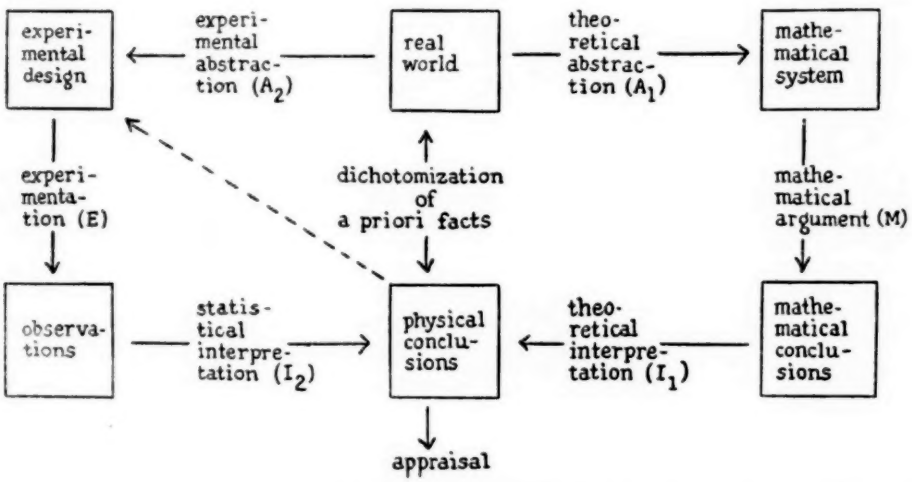

FIG. 2. A generalization of Fig. 1

left and right sides of Fig. 1 and 2. There is no natural or necessary order in which these routes should be followed. The history of science is replete with instances in which physical experiments have suggested axiom systems to the mathematicians and, thereby, contributed to the development of mathematics. On the other hand, mathematical systems developed under such stimulation, in turn, have suggested experiments. And there have been many instances of mathematical systems developed without reference to any known reality which subsequently filled a need of theoretical scientists. The direction that mathematics has taken is in considerable part due to its interaction with the physical sciences and the problems arising therein.

It is illuminating here to observe the way in which the models of the mathematical theory of probability and statistics fit this picture. As in any abstract system, the mathematical theory of probability is devoid of any real-world content; and as in any other mathematical system, the axioms of probability specify interrelationships among undefined terms. It is common to let the notion of probability itself be undefined and to attempt to capture in the axiomatic structure properties of probability motivated by the interpretations we have in mind (e.g., gambling games, processes of physical diffusion, etc.). Given an association of probabilities to prescribed elementary sets, the axioms of probability dictate how one must associate probabilities with other sets. How we make these preliminary associations, provided that we have consistency, is not relevant to the purely abstract system. When we come to apply the probability model, we are confronted with the problem of identifying real events with abstract sets in the mathematical system and the measurement problem of associating probabilities to these abstract sets. Experience has taught us that if we exploit the notion of the relative frequency of occurrence of real events when making our preliminary associations, then the interpretations from the model have a similar frequency interpretation in the real world. To be sure, our rules of composition in the formal system were devised with this in mind. We associate probabilities in one way rather than another in the process A so that when we generate AMI, our interpretations are in "close" accord with results of experimentation, T, when T is possible. When T is not possible we have to rely to a great extent on analogy.

An extremely important problem of statistics can be viewed as follows: For a priori reasons we may have a well-defined family of possible probability associations. Each element of this family, when used in the abstraction process A, generates by AMI a probability measure having a frequency interpretation over real events. In addition, we are given a set of possible actions to be taken. Preferences for these actions depend in some way on the relative "appropriateness" of different probability associations in the abstraction process, A. By conducting an experiment, T, and noting its outcome we gain some insight into the relative "appropriateness" of the different probability associations and thus base our action accordingly. Variations of this problem, which involves the entire AMI-T process, have been abstracted sufficiently so that models of mathematical statistics include counterparts of all these ingredients within the mathematical system itself.

In a given model we may be confronted with the problem of deciding

whether the AMI argument gives results "close enough" to the experimental results from T. We often can view this problem involving a complete AMI-T process as the real-world phenomenon to which we apply the A process, sending it into a formal mathematical statistics system. The statistics system analyzes step M, and our interpretation I takes the form of a statement of acceptance or rejection concerning the original theory.

The process of measurement, corresponding to A in Fig. 1, provides an excellent illustration of the role of mathematical models. There are many types of observations that can be called "measurement." Perhaps the most obvious are those made with yardsticks, thermometers, and other instruments, which result immediately in the assignment of a real number to the object being measured. In other cases, such as the number of correct items on a mental test or the size of a herd of cattle, the result of measurement is a natural number (positive integer). In still other cases, such as relative ability of two chess players, relative desirability of a pair of pictures, or relative hardness of two substances, the result is a dominance (or preference) relation. We might even stretch the concept of measurement to include such processes as *naming* each element of some class of objects, or the photographic representation of some event, or the categorization of mental illnesses or occupations.

The process of measurement may be described formally as follows. Let $P = \{p_1, p_2, \cdots\}$ denote a set of physical objects or events. By a measurement A on P we mean a function which assigns to each element p of P an element $b = A(p)$ in some mathematical system $B = \{b_1, \cdots\}$. That is, to each element of P, we associate an element of some abstract system B (the process A of Fig. 1). The system B consists of a set of elements with some mathematical structure imposed on its elements. The actual mapping into the abstract space B comprises the operation of measurement. The mathematical structure of the system B belongs to the formal side of measurement theory. The structure of B is dictated by a set of rules or axioms which states relationships between the elements of B. However, no connotation can be given to these elements of B unless it is explicitly stated in the axioms; i.e., their labels are extraneous with respect to considerations of the structure of B.

After making the mapping from P into B, then one may operate with the image elements in B (always abiding by the axioms, process M of Fig. 1). Purely mathematical results obtained in B must then be interpreted back in the real world (the process I of Fig. 1) to enable one to make predictions or to synthesize data concerning set P.

If the manifestations of P (as a result of the process T of Fig. 1) are in conflict with the results of process I obtained from B, then one must search for a new cycle AMI. Suppose that we have a family of abstractions $\{A_\alpha\}$ from the given situation P, and suppose that M_α, I_α complete the cycle begun with A_α. Among all of the available cycles $A_\alpha M_\alpha I_\alpha$ we seek one, say $A_0 M_0 I_0$, which is "closest" to T according to some criterion. Some models have a criterion built in to judge closeness, and others of a more deterministic nature require an exact fit.

The process T represents the experimental or operational part of model building; the process M represents the formal or logical aspect. The processes A and I are really the keys to the model and serve as bridges between experiment and formal reasoning.

It might be well here to draw clearly the distinction between a model and a theory. A model is not itself a theory; it is only an available or possible or potential theory until a segment of the real world has been mapped into it. Then the model becomes a theory about the real world. As a theory, it can be accepted or rejected on the basis of how well it works. As a model, it can be right or wrong only on logical grounds. A model must satisfy only internal criteria; a theory must satisfy external criteria as well.

An example of the distinction between models and theories lies in the domain of measurement. A measurement scale, such as an ordinal, interval, or ratio scale, is a model and needs only to be internally consistent. As soon as behavior or data are "measured" by being mapped into one of these scales, then the model becomes a theory about those data and may be right or wrong. Scales of measurement are only a very small portion of the many formal systems in mathematics which might serve as image spaces or models, but will be discussed here as they constitute very simple and immediate examples of the role of mathematical models. First to be discussed will be the models of conventional measurement theory, and then a generalization of these models will be presented.

MATHEMATICAL MODELS OF CLASSICAL MEASUREMENT THEORY

The first comprehensive classification of the mathematical models used in conventional measurement theory was made by Stevens (**9**). He classified scales of measurement into nominal, ordinal, interval, and ratio scales, the latter two christened by him. A more complete discussion of these scales is contained in a later work by him (**10**), and also in Coombs (**3, 5**) and Weitzenhoffer (**11**). Because of the available literature on these scales and because they constitute a restricted class of models, they will be briefly summarized here only to provide a basis for generalization in the next section.

The mathematical model of measurement is said to be *nominal* if it merely contributes a mapping A_0 of P into M_0 without imposing any further structure on M_0. A nominal scale M may be subjected to any 1–1 transformation without gain or loss in information.

An *ordinal* scale of measurement is implied if there is a natural ranking of the objects of measurement according to some attribute. More precisely, the ordinal scale is appropriate if the objects of measurement can be partitioned into classes in such a manner that (*a*) elements which belong to the same class can be considered equivalent relative to the attribute in question; (*b*) a comparative judgment or an order relation can be made between each pair of distinct classes (for example, class x is more ____ than class y); (*c*) there is an element of consistency in these comparative judgments —namely, if class x is more ____ than class y and class y is more ____ than class z, then class x is more ____ than class z (that is, the comparative judgment or order relation is transitive). For example, the familiar socioeconomic classes, upper-upper, lower-upper, upper-middle, lower-middle, upper-lower, and lower-lower, imply the measurement of socioeconomic status on an ordinal scale. The numbers 1, 2, 3, 4, 5, 6, or 1, 5, 10, 11, 12, 14, or the letters A, B, C, D, E, F could designate the six classes without gain or loss of information.

The measurement is said to be an *interval* scale when the set M consists of the real numbers and any linear

transformation, $y = ax + b$ $(a \neq 0)$, on M is permissible. Measurement on an interval scale is achieved with a constant unit of measurement and an arbitrary zero. An example of an interval scale is the measure of time. That is, "physical events" can be mapped into the real numbers and all the operations of arithmetic are permissible on the differences between all pairs of these numbers.

If the set M consists of the real numbers subject only to the transformation group $y = cx$ where c is any nonzero scalar, the scale is called a *ratio* scale. Measurement on a ratio scale is achieved with an absolute zero and a constant unit of measurement. The scalar c signifies that only the unit of measurement is arbitrary. In a ratio scale all the operations of arithmetic are permissible. The most familiar examples of ratio scales are observed in physics in such measurements as length, weight, and absolute temperature.

A Generalization of Measurement Models

An axiomatic basis for certain scales of measurement will be presented in this section. Other scales can be generated by forming mixtures (or *composites*) of these. Indeed, some of the scales listed in the diagram shown in Fig. 3 can be regarded as composites of others.

We will now list defining axioms for each of these systems and briefly discuss their roles. It is not claimed that this list is exhaustive; it is presented to illustrate certain possibilities for significant generalizations of scales used in the classical theory. The arrangement in the diagram is from top to bottom in order of increasing strength of axioms; a connecting line indicates that the lower listed system is a special case of the higher one.

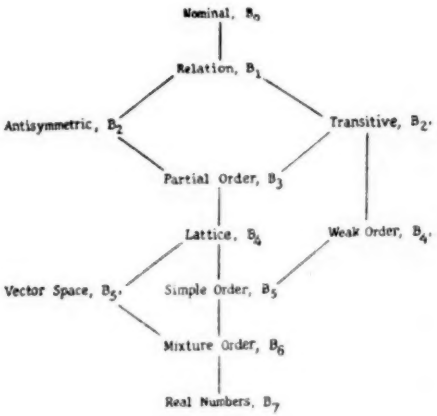

FIG. 3. Measurement scales

B_0, *the nominal scale.* A nominal scale, B_0, may be considered a mathematical system consisting merely of a set of elements. We define the *index* of B_0 to be the number of elements in B_0. (The index may be finite or infinite.)

Examples of segments of the real world that are mapped into nominal scales are psychiatric classifications, job families, and disease types.

The nominal scale, B_0, is the most primitive step in any system of measurement. The process of naming partitions a set into classes such that there is a relation of "equality" or equivalence between pairs of elements from the same class. The nominal scale is fundamental since the process of discrimination is a necessary prerequisite for any more complex form of measurement.

B_1, *the relation scale.* Perhaps the smallest step that may be taken to strengthen this mathematical system is to introduce a relation between some pairs of elements. In technical language, a relation R on a set B_1 is a set of ordered pairs (b, b') of elements of B_1. We write bRb' to indicate that (b, b') is one of the pairs included in the relation R, and call the set B_1 a relation scale. It is important to rec-

ognize that for R to constitute a very useful relation, not all possible pairs (b, b') from B_1 can be included in the relation R.

With some risk of misinterpretation or distortion, these concepts might be illustrated as follows. Consider a set of persons identified by a nominal scale, B_0. Let us now define the relation R on B_0 to be "loves." Thus R consists of the ordered pairs (a, b) for which, a loves b.

The particular relation used here as an illustration is one whose mathematical properties are mostly negative. We cannot conclude from a loves b and b loves c that a loves c, or that b loves a, or that b does not love a. For example, if John loves Mary and if Mary loves Peter, it may well be that, far from loving him, John would like to see Peter transported to the South Pole. In the terminology to be introduced below, we would say that love is not symmetric, is not asymmetric, and is not transitive.

B_2, *the antisymmetric relation scale.* A relation R on a set B is said to be *antisymmetric* if aRb and bRa together imply that a is identical with b. An example is the relation $\geqq$ for real numbers. A statement such as "picture a is at least as good as picture b" illustrates an antisymmetric relation on a collection of pictures, provided that there are not in the collection two distinct pictures of equal merit, i.e., two pictures about which the judge is *indifferent.*

Closely connected to the concept of antisymmetry is that of asymmetry. A relation R on a set B is said to be *asymmetric* if aRb implies $bR'a$ (where $bR'a$ means that b is *not* in the relation R to a). The mathematical prototype of asymmetry is the relation $>$ for real numbers. Verbal forms for asymmetric relations include such statements as "picture a is better than picture b," or "player a beats player b in a game."

Antisymmetry and asymmetry are seen to be at the root of statements of comparison. These two classes of relations can be regarded as the most primitive types of order relations. At the opposite pole from these concepts is that of symmetry. A relation R is said to be *symmetric* if aRb implies bRa. For example, the relations "is a sibling of," "is a cousin of," and "is the same color as" are all symmetric.

If S is an asymmetric relation on a set B, we can obtain from it an antisymmetric relation R by the definition: aRb means either aSb or $a = b$. Conversely, if R is antisymmetric and we define aSb to mean aRb and $a \neq b$, then S is asymmetric. It is customary to use the symbols $\leqq$, $\geqq$ for antisymmetric relations and to use $<$, $>$ for the associated asymmetric relations.

$B_{2'}$, *the transitive relation scale.* A relation R is said to be transitive if aRb and bRc imply aRc. In the physical world, preference judgments which are not transitive are frequently regarded as inconsistent or irrational. However, situations such as that of three chess players, each of whom can beat one of the other two, show that transitivity is not a requirement of nature.

The chess player relation is antisymmetric but not transitive. An example of a relation that is symmetric and transitive is given by a communication system where each link is bidirectional; here aRb is given the meaning "there exists a chain of links starting with a and ending with b." If the links are not required to be bidirectional, the relation is still transitive but is no longer symmetric. Note that in this example it is quite possible to have aRa, i.e., a chain be-

ginning at a and ending at a. (This chain must have at least one element different from a.)

The relation "a is the rival of b" (say as suitors of a particular girl) is symmetric and is almost transitive. If aRb and bRc, we can conclude aRc unless $a = c$; but we can hardly regard a as being his own rival. This type of relation arises frequently in studies of social structures. We say a relation R is *quasi-transitive* if aRb, bRc, and $a \neq c$ imply aRc. The sibling relation is also quasi-transitive. Of course, if R is quasi-transitive we can define a new relation S to be the same as R except that also aRb, bRa imply aSa. In some instances S is just as good a model as R, but in others the extension from R to S destroys the usefulness of the model.

As an example consider the structure matrix $A = \|a_{ij}\|$ of some society. Thus we set $a_{ij} = 1$ if person i has direct influence on person j and set $a_{ij} = 0$ otherwise. One must decide in accordance with the purpose of the investigation whether or not to set the diagonal element a_{ii} equal to 0 or to 1. (The relation "i has direct influence on j" is not transitive even if we take each $a_{ii} = 1$, but this example nevertheless illustrates the kind of problem involved in the contrast between transitivity and quasi-transitivity.)

If the relation aRb meant a "is higher in socioeconomic status than" b, and this required that a had more income *and* more education than b, then the relation R would be asymmetric and transitive.

B_3, *the partly ordered scale.* A relation $\geqq$ which is reflexive,[5] antisymmetric, and transitive is called a partial order. If for some pair (a, b) neither of the relations $a \geqq b$ nor $b \geqq a$ holds, we say that a and b are *incomparable* relative to $\geqq$. In the case of a preference relation, incomparability is not the same thing as *indifference.* We call a set a *poset* (*p*artly *o*rdered *set*) if there is a partial order relation defined on B.

If $a \geqq b$, we also write $b \leqq a$; if $a \geqq b$ and $a \neq b$, we write $a > b$ or $b < a$.

A partial order may be illustrated as follows. Suppose that on a mental test no two individuals in a group pass exactly the same items. Now let $a \geqq b$ symbolize the relation "a passed every item b did and perhaps more." Then $a > b$ means that "a passed all the items b did and at least one more." This poset reflects multidimensionality of the attributes mediating the test performance, and some interesting mathematical problems arise regarding the partial order as a "product" of simple orders. The result is a nonmetric form of factor analysis with some of the same problems as factor analysis (**3**).

Next we consider the mental test example modified so as to allow the possibility that two individuals a and b pass exactly the same items. Then in the above notation we have $a \geqq b$ and $b \geqq a$, but not $b = a$. Hence, $\geqq$ no longer gives a partial order. However, if we define $a \Im b$ to mean $a \geqq b$ and $b \geqq a$, it is not hard to show that if we identify individuals with the same test performance then $\geqq$ is a partial order relation. Or, alternatively, we could consider $\geqq$ as a partial order relation on the set of possible test performances. It is customary to make such identifications and speak of a partial order as if it were actually on the initial set rather than on the identified classes or on the test results.

[5] A relation is reflexive if it holds between an element and the same element, symbolized aRa. For example, the relation $\geqq$ on numbers is reflexive and the relation "in the same family as" is reflexive.

Another example of a partial order is implicit in the treatment of the comparative efficiency of mental tests on a "cost-utility" basis (**1**). "Cost" is the fraction of potentially successful people who are eliminated by a test; "utility" is the fraction of potential failures who are eliminated by the test. If for their respective cutting scores one test has a higher utility and a lower cost than another it is a superior test, but if it had a higher utility *and* a higher cost the two tests would be incomparable unless the relative weight of excluding a potential success to including a potential failure were known.

A basic problem in the theory of testing hypotheses in statistical inference is to test a simple hypothesis, H_0 (null hypothesis), against a single alternative hypothesis, H_1, by means of experimental data. A test, T, associates to each experimental outcome the decision to accept H_0 or to accept H_1 (but not both!). Each test T is appraised by a pair of numbers, namely, the probability of accepting H_1 if H_0 is true, $P_T(H_1|H_0)$, and the probability of accepting H_0 if H_1 is true, $P_T(H_0|H_1)$. Given two tests, T' and T'', then T' is said to be as good as T'' ($T' \geqq T''$) if and only if

$$P_{T'}(H_1|H_0) \leqq P_{T''}(H_1|H_0)$$
$$P_{T'}(H_0|H_1) \leqq P_{T''}(H_0|H_1)$$

The relation $\geqq$ on the set of all tests is an example of a partial order.

B_4, *lattice.* Let B be a poset relative to a relation $\geqq$. If a, b, c are elements of B and $c \geqq a$, $c \geqq b$, we say that c is an *upper bound* of a and b. If also $c \leqq x$ for every upper bound x of a and b, we say that c is the *least upper bound* of a and b and write $c = a \cup b$. In terms of the example of mental testing, c could be a person who passed exactly those items which were passed by at least one of a and b. Analogously, if $d \leqq a$, $d \leqq b$, we say that d is a *lower bound* of a and b, and if also $d \geqq y$ for all lower bounds y of a and b, we say that d is the *greatest lower bound* of a and b and write $d = a \cap b$. In our example d could be a person who passed exactly those items passed by both a and b.

A pair a, b need not have a least upper bound or a greatest lower bound. For example, if a passed items 1, 2, 3, 4; b passed 1, 2, 5, 6; c passed 1, 2, 3, 4, 5, 6, 7; c' passed 1, 2, 3, 4, 5, 6, 8; and there are no other persons, then both c and c' are upper bounds to a and b but there is no least upper bound. Also in this case there are no lower bounds for a and b and hence no greatest lower bound.

A poset is said to be a *lattice* if, for every pair a, b, both $a \cup b$ and $a \cap b$ exist. The lattice is an intermediate model between a partial order and a vector space.

George Miller[6] (Massachusetts Institute of Technology) has recently investigated the use of a lattice theoretic treatment of information in experimental psychology. To each item of information he associates (process A) an element of a lattice. If two items of information are associated respectively with elements x and y of the lattice, then the item which consists of the information common to the original items is associated with the element $x \cap y$, and the item which consists of the information contained in either of the original items is associated with the element $x \cup y$. As used by Miller, an item of information might consist of a cue or a sequence of cues in an experimental situation. The common procedure is to summarize the structure of the experiment by means of a lattice, given an experimental setup. However, the

[6] Personal communication.

abstract lattice in turn can motivate new types of experimental situations and indicate analogies between experimental designs which otherwise would not be apparent.

$B_{4'}$, *weak order.* A transitive order $\leqq$ is defined on $B_{4'}$ and has the property that for every pair a, b either $a \leqq b$ or $b \leqq a$. If both $a \leqq b$ and $b \leqq a$, we say that a and b are indifferent. Indifference is an equivalence relation (i.e., is reflexive, symmetric, and transitive).

A weak ordering would be illustrated by the military ranks of second lieutenant, first lieutenant, captain, major, etc. Each of these would constitute an equivalence class, and for any two officers (a, b), either $a \geqq b$ or $b \geqq a$, or both.

B_5, *chain.* A poset in which every pair is comparable is called a chain (or simple order, or linear order, or complete order). Alternatively, a chain is a weak order in which each indifference class consists of a single element. Here every pair of elements is ordered.

The previous example of a weak ordering of military rank could be converted into a chain if date of rank, standing in class, etc. were taken into account. Then, for every two distinct elements, a, b, either $a > b$ or $b > a$.

The ordinal scales of classical measurement theory are examples of chains.

$B_{5'}$, *partly ordered vector space.* A special case of lattice is provided by a real vector space (or a subset of a vector space). A *vector* $x = (x_1, \cdots, x_n)$ is an ordered set of n real numbers called the *components* of the vector. We define $x \leqq y$ to mean that $x_i \leqq y_i$ for each component. (Here the second symbol $\leqq$ refers to the usual ordering of real numbers.) This definition makes the vector space into a poset, and this poset is a lattice which is called a partly ordered vector space.

A partly ordered vector space is illustrated by the comparability of individuals in mental abilities. Conceiving of intelligence as made up of a number of primary mental abilities, each of these constitutes a component or dimension. Then, it may be said of two individuals x and y that y is at least as intelligent as x if and only if y has as much or more of each component as x has.

The term *vector* is sometimes used in a more general situation. If $C_1, \cdots, C_n$ are chain orders we may consider vectors or n-tuples $c = (c_1, \cdots, c_n)$ where the ith component c_i lies in the chain order C_i $(i = 1, \cdots, n)$. The set C of all such vectors c is called the *Cartesian product* of $C_1, \cdots, C_n$ and is denoted by $C = C_1 \times \cdots \times C_n$. We can make C into a poset by a process analogous to that used above for real vector spaces. Note that a real vector space is the special case of a Cartesian product of n factors $C_1, \cdots, C_n$ each equal to the set of real numbers.

B_6, *simply ordered vector space, or utility space.* A real vector space (or subset) in which $x < y$ is defined lexicographically, i.e., $x < y$ if $x_1 = y_1, \cdots, x_{i-1} = y_{i-1}, x_i < y_i$ is a special case of simple order.

A lexicographic ordering can be illustrated by the manner in which we might expect a fortune hunter to simply order a number of unmarried women. Presumably, financial assets would be the principal component and he might construct a weak ordering of the women, into, say five classes, on this basis. Then he would turn to the second component, say beauty, and within each of the financial classes construct a simple ordering of the women on this component. Any two women (a, b) would then be simply ordered as follows:

1. If a were in a higher financial class than b, a would be preferred to b.

2. If a were in the same financial class as b, then preference would be determined by their relation on the beauty component.

B_7, *real numbers.* The transition to scales using the real numbers has been given additional importance by the development of von Neumann-Morgenstern utilities. Even though one may wish to arrive here in order to have a simple index when a decision is to be made, it may frequently be desirable not to get here all at once, but to keep the components at a weaker level until it is necessary to map into the real numbers.

Measurement scales involving the real numbers, the interval scale, and the ratio scale have been discussed at length in the literature (**3**, **5**, **9**, **10**, **11**) and will not be pursued again here.

FURTHER EXTENSIONS

The various mathematical systems discussed here as available for measurement have been illustrated with objects of the real world mapped into the elements of an abstract system. A further level of abstraction is provided by defining a "distance function" in the abstract system, in which ordered pairs of elements in the abstract system are mapped into elements of another abstract system about which a variety of assertions may be made. In the context of measurement, these pairs of elements may correspond to "differences" between pairs of objects in the real world. These differences may themselves then be mapped into an appropriate abstract system such as one of those discussed here.

A number of these types of scales have been discussed by one of the authors (**4**, ch. 1). An illustration is the ordered metric scale in which the objects themselves satisfy B_5, a simply ordered scale, and ordered pairs of objects, regarded as "distances" between them, satisfy B_3, a partly ordered scale. Such scales are now being utilized for the measurement of utility and psychological probability in experiments on decision making under uncertainty (**6**, **7**).

SUMMARY

One role of mathematical models is to provide a logical route to go from characteristics of the real world to predictions about it. The alternative route is by observation or experiment on the real world itself. The view expressed here is that these two routes are coordinate.

The various scales used in measurement serve as an illustration of the application of mathematical models and are subject to the same constraints as other mathematical models. That is, if the axioms underlying the scale are not satisfied by that segment of the real world which is mapped into it, then the interpretations of the mathematical conclusions may have no reality or meaning. Thus, to insist that measurement always constitutes the mapping of physical objects into the real number system is to impose on the real world an abstract theory which may be invalid.

A partial ordering of various alternative mathematical systems available for measurement has been presented with illustrations in order to reveal the relative strengths of these scales to which the real world must conform to permit their application. We make no claim to completeness in this list of models for measurement theory. Our purpose is to point out the richness of the set of possible models and to give some examples that show how the use of more general models can extend the domain of classical measurement theory.

None of the discussion here should be taken as an argument for the use of weaker scales in the place of stronger scales for their own sake. The measurement scale utilized constitutes a theory about the real world and the stronger the theory the better, so long as it is correct. The addition to a scale of axioms which are not satisfied by the real world is a step away from the path of progress.

REFERENCES

1. BERKSON, J. "Cost-utility" as a measure of the efficiency of a test. *J. Amer. stat. Ass.*, 1947, **42**, 246–255.
2. BIRKHOFF, G. *Lattice theory*. New York: American Mathematical Society, 1948.
3. COOMBS, C. H. Mathematical models in psychological scaling. *J. Amer. stat. Ass.*, 1951, **46**, 480–489.
4. COOMBS, C. H. A theory of psychological scaling. *Engng Res. Bull.* (Univer. of Michigan), 1952, No. 34.
5. COOMBS, C. H. The theory and methods of social measurement. In L. Festinger & D. Katz (Eds.), *Research methods in the behavioral sciences*. New York: Dryden Press, 1953. Pp. 471–535.
6. COOMBS, C. H. Social choice and strength of preference. In C. H. Coombs, R. M. Thrall, & R. L. Davis (Eds.), *Decision processes*. New York: Wiley, in press.
7. COOMBS, C. H., & BEARDSLEE, D. C. On decision making under uncertainty. In C. H. Coombs, R. M. Thrall, & R. L. Davis (Eds.), *Decision processes*. New York: Wiley, in press.
8. KERSHNER, R. B., & WILCOX, L. R. *The anatomy of mathematics*. New York: Ronald, 1950.
9. STEVENS, S. S. On the theory of scales of measurement. *Science*, 1946, **103**, 677–680.
10. STEVENS, S. S. Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1950. Pp. 1–49.
11. WEITZENHOFFER, A. M. Mathematical structures and psychological measurements. *Psychometrika*, 1951, **16**, 387–406.
12. WEYL, H. *Philosophy of mathematics and natural science*. Princeton: Princeton Univer. Press, 1949.
13. WILDER, R. L. *Introduction to the foundations of mathematics*. New York: Wiley, 1952.

(Received June 1, 1953)